DOCUMENT RESUME

ED 166 192 95 SP 013 890

AUTHOR Evertson, Carolyn M.; And Others
TITLE Process-Outcome Relationships in the Texas Junior High School Study: Compendium.
SPONS AGENCY National Inst. of Education (DHEW), Washington, D.C.
REPORT NO RD-4064
PUB DATE Mar 78
CONTRACT 400-76-0073; NIE-G-78-0116
NOTE 93p.; Papers presented at the Annual Meeting of the American Educational Research Association (Toronto, Canada, March 1978) ; Pages with light type may not reproduce clearly

EDRS PRICE MF-$0.83 HC-$4.67 Plus Postage.
DESCRIPTORS *Academic Achievement; Class Management; Classroom Observation Techniques; Classroom Research; *Effective Teaching; English Instruction; Mathematics Instruction; *Measurement Techniques; Rating Scales; *Statistical Analysis; Student Evaluation of Teacher Performance; *Teacher Behavior; Teacher Characteristics; Teacher Evaluation; Teaching Techniques

ABSTRACT

This report presents the process-outcome relationships discovered in the Texas Junior High School Study of correlates of effective teaching. Classroom processes and cognitive and affective indices of outcome are discussed for two subject areas (English and math) and for classrooms with varying levels of entering ability. The four separate papers in this compendium focus on the following : (1) overview of methodology and rationale; (2) high inference ratings of general behaviors; (3) low inference coding of specific behaviors; and (4) comparison of findings for cognitive and affective criteria. The results indicated that student characteristics were most influential on learning outcomes in English classes, but that teacher behavior was more important in math classes. Effective instruction was more readily identifiable in the math classes as well. Further results emphasized that caution should be exercised in any attempt to use student ratings in teacher evaluation. The data for the research is presented in tables, and references are provided. (DS)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

ED166192

U.S. DEPARTMENT OF HEALTH,
EDUCATION & WELFARE
NATIONAL INSTITUTE OF
EDUCATION

THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL NATIONAL INSTITUTE OF EDUCATION POSITION OR POLICY

Process-outcome Relationships in the Texas Junior High School Study: Compendium

Carolyn M. Evertson

Linda M. Anderson

R&D Center for Teacher Education
The University of Texas at Austin

and

Jere E. Brophy

Michigan State University

SP 013 890

R&D Report No. 4064

Process-Outcome Relationships in the Texas Junior High School Study:

Overview of Methodology and Rationale

Carolyn M. Evertson

Linda M. Anderson

R&D Center for Teacher Education
The University of Texas at Austin

and

Jere E. Brophy

Michigan State University

R&D Report No. 4064

(COET No. 78-8.1)

This report is part of a four paper session entitled "Process-Outcome Relationships in the Texas Junior High School Study," which was presented at the Annual Meeting of the American Educational Research Association, Toronto, Canada, March, 1978. This paper is a general summary of part of the data from the study and should not be cited without further reference to the full report.

This study was supported in part by the National Institute of Education Contract OB-NIE-G-78-0116, The Research and Development Center for Teacher Education. Dr. Brophy is also supported by NIE Contract 400-76-0073, The Institute for Research on Teaching. The opinions expressed herein do not necessarily reflect the position or policy of the National Institute of Education, and no official endorsement by that office should be inferred.

Introduction

This report presents the process-outcome relationships found in the data from the Texas Junior High School Study, conducted by the Correlates of Effective Teaching Program at the University of Texas Research and Development Center for Teacher Education. Both low and high inference measures of classroom processes and cognitive and affective indices of outcome will be discussed for two subject areas and for classrooms with varying levels of entering ability. Results describe patterns of teaching behaviors which relate to either type of outcome in different contexts, and may therefore suggest what composes "effective teaching" at the junior high school level. Although there has been some research on this topic at the elementary level, there has been no work done on teaching in the secondary school that has included the large data base, variety of measures, and specification of contextual influences included in this study. These data were analyzed at the class level and focus only on process-outcome relationships.

Background. The Junior High School Study was conceived as a replication and extension of an earlier study of teaching effectiveness conducted at the second and third grade levels (Brophy and Evertson, 1976). The earlier study suggested several effective strategies for teaching elementary students, but it did not support several variables popular among educational researchers, such as indirect teaching, extensive use of class discussion, and pupil talk. One question arising from these results was that,

even though such strategies were not related to achievement in the early grades, would they become more important at the later grades when most "tool" skills are mastered and students are learning to apply them? Another question was, to what extent do contextual influences affect such relationships? Therefore, the Junior High School Study was designed as an effort at replication of the earlier process-outcome study but at different grade levels, and also as a more extensive examination of teaching behaviors that were related to both cognitive and affective student outcomes.

To address these objectives, the design was improved from the earlier study in several ways: 1) data were collected during the same school year in parallel sections of seventh and eighth grade mathematics or English classes taught by the same teachers in the same public schools; 2) two observers alternated visits to 136 classrooms, averaging 20 hours of observation per classroom; 3) data were collected on a large number of individual students, enabling investigations of student effects as well as teacher effects; and 4) the low inference observational coding system was modified especially for use in secondary classrooms in order to capture appropriate contextual differences.

Methodology

Selection of subjects and design of instruments reflected the important research questions of the study. In all, 136 classes in nine schools were observed. They were chosen so that:

1. Two different but important school subject areas were in-

cluded--math and English--making it possible to investigate differences in effective teaching strategies in different settings.

2. The nine junior high schools represented a wide range of socio-economic status (SES) and achievement levels, making it possible to examine differences in effective teaching strategies for low vs. high ability classes.

3. Each participating teacher was observed in two separate sections of his or her subject matter (math or English), allowing systematic attention to the question of teacher stability versus variability in process behaviors across classroom settings, as well as to the central question of teaching effectiveness as it was affected by grade, subject matter, student sex, and other context differences.

Subjects

Description of teachers in the sample. Sixty-eight teachers (39 English and 29 math) were observed in nine of the eleven junior high schools in a large urban school district. (Two other junior high schools were not included because they were using an exclusively self-paced mathematics program that allowed for very little public teacher-student interaction.) Because two sections were observed for each teacher, there were 136 classrooms in all. Two observers alternated visits to these classes, for an average of 20 one-hour observations per class, throughout the school year 1974-1975. (The actual range was from 16 to 22 observations.)

Attempts were made to avoid unique situations by excluding unusual scheduling or split sections that met for part of the time before lunch and the remainder of the time after lunch.

Teachers selected for the study were those with at least one previous year of experience in their subject matter area. Student teachers, first year teachers, or teachers who shifted into these areas from some other subject matter were not included.

The resulting teacher sample was unusually complete, and was reasonably free of volunteer effects or other sample bias effects, since nearly all the eligible faculty from each of the nine junior high schools participated.

Description of students in the sample. The nine junior high schools represented a wide range of socio-economic status and achievement levels. They were included in a local desegregation plan which provided for busing of black students only to predominantly white junior high schools. Although it was necessary to collect data on individual students, it was apparent that observers would not be able to identify and remember code numbers for all students in each class in which they observed (some observers saw as many as 500-600 students each week). Therefore, in order to be able to record at least some individual student data, a subsample of 10-12 "target students" was selected randomly, within sex, in each class. These target students (N = 1614) were selected from teachers' rolls before observations were conducted in any classrooms.

One exception to random selection was made, however. In selecting target students, efforts were made to include a large sample of students who were attending both a math and an English class included in the study (N = 199). These selection procedures resulted in a subgroup of students who were taught by two different teachers and are referred to as "overlap students." Therefore, the design

of the study made it possible to examine not only stability of individual teachers' behaviors across sections, but also stability of individual student behavior across teachers and subjects, as well as comparisons of appropriate student and teacher behavior in terms of learning gains in different subject matters. A thorough discussion of these stability findings may be found in Evertson et al. (Note 1).

Instruments

A wide variety of instruments was used to collect data in this study. They can be broadly classified as either process measures, which described classroom occurrences, or outcome measures, which described the achievement and attitudes of the students at the end of the year. The results presented in this report are based on relationships between each of the process measures and each outcome.

Description of process measures. The process instruments can be classified as being low inference or high inference. The former was an observation system used to note the frequency of occurrence of several discrete behaviors, and the latter consisted of two kinds of rating scales. Some were completed during the year and then averaged, and some were completed one time at the end of the year.

Low inference process measures: The Classroom Observation Coding System. This was an adaptation of the coding system used in the Texas Teacher Effectiveness Study (Brophy and Evertson, Note 2; Brophy, Evertson, Baum, Crawford, and Edgar, Note 3). The modified instrument was developed to include a wide range of variables, including those used most frequently in previous educational research, as well as some unique to this study. The major adaptations and expansions

were done to add variables based on Kounin's (1970) research on classroom management techniques, and to break down teacher behavior more specifically according to context variables having to do with the time and nature of classroom interaction during which a particular observation took place.

For example, while using the coding system, observers recorded the amounts of time teachers spent in various activities, such as class discussion, drill, lost time, transitions, etc. They also noted the context area of the lessons for that day (e.g., division with whole numbers or fractions for math classes, or grammar, drama presentations, literature, etc., for English classes). Such information was useful for placing frequency data within the appropriate context.

Another addition to the coding system was provision for a detailed recording of student misbehaviors (e.g., mild misbehaviors, socializing, sassing, verbal or physical aggression) and the manner in which the teacher handled the incident. In addition, observers recorded the appropriateness of the disciplinary intervention according to categories suggested by Kounin (target error, timing error, overreact, ignore). This allowed examination of not only the type of student misbehavior, but also teacher reaction to it and its appropriateness.

In all, the system was more complex and detailed than previous systems (Brophy and Evertson, Note 2) so as to allow recording of behaviors which were more likely to occur with older students. The system was expanded to include categories allowing more detailed coding of teacher-initiated versus student-initiated public response

opportunities, private contacts initiated by student or teacher (work-related, procedural, or personal-social) and classroom behavior-related incidents.

Observers were trained to a reliability criterion of 80% agreement on each major section of the system, computed according to the following formula:

$$\% \text{ agreement} = \frac{\text{Codes agreed upon by Coders A \& B}}{\text{Coder A's codes (which Coder B missed) + Coder B's codes (which Coder A missed) + those coded and agreed on by both, + those coded by both but disagreed on}}$$

See Coulter (Note 4) for a detailed explanation of training procedures.

After all observations were completed, the low inference data were reduced. The coding system yielded 768 frequencies which were tallied over all observations for a single class. These consisted of sums of single categories (e.g., the sum of correct answers), sums of two or more categories applicable to the same interaction (e.g., the sum of all correct answers receiving praise), and aggregates of single categories over many interactions (e.g., the sum of all answers given by the students).

These frequencies were computed for each of the 136 observed classes, and were then used to create other scores which were more appropriate for analysis.

These "final-form" variables were of two kinds: (1) rate scores, for which frequencies were divided by number of minutes per average class period (50), thus giving an index of the mean absolute rate at which certain behaviors occurred (e.g., "correct answers per observation"), and (2) proportion scores, which were computed by

dividing raw frequencies of the variables in the coding system by the frequencies of the major categories, in order to see the relative occurrence of given behaviors. For example, the "proportion of process questions" was computed by dividing frequency of these questions by the total for all questioning categories; therefore, the proportions of each of the four question types (process, product, choice, and opinion) add to 100%.

Some of the proportion variables were composed of frequencies describing the simultaneous occurrence of two discrete categories in the coding system. For example, the measure "proportion of student behaviors with management and no error" reflects the proportion of behavior contacts coded as management response (vs. nonverbal intervention, criticism, or threat) and as containing no error (vs. a target error, a timing error, or an overreaction). Each behavioral contact that was solved with only management response, and solved in a way that involved no error, counted toward the total used in the numerator of this proportion. The sum of these behavioral contacts was divided by the total number of behavioral interventions observed.

Some of these proportion measures involved more than one frequency score in the numerator or denominator. For example, the variable "proportion of don't know or no response answers after which teacher gave the answer" included both the frequency of "don't know" and of "no response" answers in the denominator. (These were combined into one variable because both were low frequency variables compared to correct and incorrect answers.) Thus, the variable "proportion of don't know and no response answers after which the teacher gave the answer" was derived by summing the times that teachers gave the an-

swer to students when they either said that they did not know or made no response, and dividing this total by the total number of times that students in the class said that they did not know or made no response.

High inference process measures: Rating scales done throughout the year. Two high inference measures were completed during each observation, and then used to calculate single scores representing averages over the year. The Classroom Observation Scales were 12 five-point scales that described certain global classroom or teacher characteristics, such as level of student attention, clarity of presentation, enthusiasm, and affect. A complete description of the scales and instructions for their use may be found in Emmer (Note 4).

Emmer reported between-observer reliability estimates (intraclass correlations) ranging from .48 to .89 for scales recording adequate variation among teachers. All 12 scales in the present study elicited such variation. The reliability of the high inference ratings is also reflected in the very strong correlations between ratings of the same teachers in their two classes observed (Evertson et al., Note 1).

Also during each observation, after completing the Classroom Observation Scales, observers rated the presence or absence of certain types of teacher questioning: memory-fact related, higher cognitive level, or personal-self questions.

After each observation of a class, observers completed Classroom Descriptions, by recording any impressions, comments and anecdotes about what occurred during the class hour. The form and focus of these class descriptions were left relatively unstructured, because investigators were interested in capturing any extra information that

was not elicited by the behavioral coding system or the observation scales. This method allowed observers to note qualitative and contextual elements of the classroom environment as well as the sequence and content of instruction. The descriptions proved invaluable for cross-checking the observation sheets during data processing, and they added an important dimension to our data on classroom events. They were scored by using a system adapted partially from that used by Tikunoff, Berliner, and Rist (Note 5), supplemented by other categories suggested by events that appeared in the narratives. Each set of protocols was scored on 34 five-point scales for such categories as reacting to students' feelings, dividing time and attention equally among students, and perceiving learning rates of students and adjusting pace accordingly. Pairs of observers first scored all protocols independently and then resolved disagreements by discussion. Sometimes the resolution involved redefining the categories in more specific ways and rescoring the descriptions.

High inference process measures: Rating scales completed at the end of the year. There were three types of ratings done at the end of the year, two by the observers, and one by teachers about their target students. The Observer Ratings of Teachers included 79 five-point scales of such attributes as personal-social interactive style, competency in subject area, and classroom organization and control.

Since each teacher was scored by more than one observer, ratings were correlated to get reliability estimates. Fifteen of the items were dropped for unreliability when $p \geq .10$. There is reason to suspect

halo effect in these ratings, because 42 teachers were seen in both of their sections by the same two observers, 22 teachers were rated by three observers, and four teachers were rated by four different observers.

The Observer Ratings of Students were also completed at the end of the year. Observers completed 26 five-point rating scales on each target student. These included characteristics such as work habits, likeability, classroom conduct, and physical development. Again, each target student was seen by at least two observers. One rating scale was dropped for unreliability, although reliability estimates (correlations) for the remaining items were high ($p \leq .01$).

Teacher Ratings of Students. At the end of the year each teacher provided ratings on five-point scales of each target student. These ratings indicate the teachers' perceptions of students' relative achievement level, motivation, work habits, likeability and classroom behavior. These ratings were also summed and averaged to yield a score for each class.

Summary of process measures. Data will be presented in this report on seven measures which describe classroom processes:

Low inference measures:

1. Rates computed from frequencies derived from coding system data
2. Proportions computed from frequencies derived from coding system data

High inference measures:

3. Classroom Observation Scales, including the present-

absent ratings of question type (completed during every observation and averaged for the year)

4. Classroom Descriptions (completed during every observation, summarized, and scored for the year)
5. Observer Ratings of Teachers (completed at the end of the year)
6. Observer Ratings of Students (completed at the end of the year)
7. Teacher Ratings of Students (completed at the end of the year)

Description of outcome measures. There are two outcome measures administered to the students at the end of the year: achievement tests in each subject area, and Student Ratings of Teachers. In addition, the students' scores on the California Achievement Test taken in the Spring of the preceding school year were used as covariables in any analyses involving the outcome measures. This combination of cognitive and attitudinal (or affective) measures was chosen in order to examine two important but different objectives that teachers might set for students in junior high school. Using these data, it is possible to examine any "trade-off" between cognitive learning and attitudes toward school and teachers that might exist.

Affective outcome measures: Student Ratings of Teachers. At the end of the school year, students were asked to fill out nine five-point rating scales about their teachers. These scales included essentially two types of items: those assessing general liking of the teacher ("I would go to this teacher if I had a problem") and those assessing the degree to which the student felt he/she learned the

subject matter ("I learned a lot from this teacher"). All students, both target and non-target, filled out these assignments. Items were factored, and one general factor emerged which was named "generalized likeability" or general liking of teacher.

The average rating for a teacher on all nine scales was used as an affective or attitudinal criterion to which all other measures could be compared.

Cognitive outcome measures: Achievement tests and CAT scores. Students' scores on the English and math subtests of the California Achievement Tests given in the Spring prior to observation were used to estimate entering ability. These subtest scores were summed, and one combined score was used as the indicator of initial ability for each student for each of the two subject matters. The summed scores were then averaged to obtain teacher scores for each class section.

To obtain an estimate of achievement at the end of the year, tests were specially constructed for use in this study to measure knowledge of English grammar, word usage, punctuation, and spelling, and to measure knowledge of mathematical computation and reasoning.

These tests, which were administered during the first weeks of May, were designed to be content valid to the extent that the items reflected the subject matter being taught in the observed classrooms. Information on the subject matter covered was gathered from observers' coding sheets. Observers were instructed to check one or more of the designated descriptors to record the content covered for the class

period. Also, observers were given copies of the tests, and they noted for each item whether or not its content was covered during their observation periods. Copies of the adopted texts were also consulted.

Prior to the administration of the tests, students were asked to fill out the student rating forms mentioned previously. These were collected, and then the achievement tests were distributed. Students were allowed approximately 45 minutes to take their respective tests. No students received perfect scores, and only a small percentage of students completed their entire tests. While each student received a single total score, the individual test items were also scored. This information was preserved so that item analyses could be performed. Items that did not discriminate were eliminated before students' aggregate scores were computed. A Rasch analysis of the tests indicated a high internal consistency in each test and also confirmed that the appropriate items had been dropped.

Decisions regarding the use of the outcome measures. The mean CAT and achievement test scores were computed to use in analyses relating process measures to cognitive outcomes. There were several important questions to answer regarding the choice of the achievement criterion and covariate. Large differences in general level of academic achievement existed among the nine schools observed in the study, and a certain amount of "tracking" was evident within schools. Also, a significant number of students lacked either the CAT score or the achievement test score, and it seemed undesirable

to exclude these students entirely from analyses. Before performing two sets of analyses with the two outcome measures, more information was needed about their relationship to one another. In addition, current controversy regarding the measurement of learning gains, especially the use of residualized gains and the need for appropriate levels of analysis, called for a careful examination of the outcome measures before pursuing the other process-outcome analyses.

Partial correlations of several possible predictor variables with achievement test scores were computed to determine the independent contribution of each predictor to the overall variance. Stepwise multiple regression analyses were performed with achievement test scores as the criterion, using the same predictor sets. Each set of analyses was done twice, once using only students with both CAT and achievement scores (restricted sample), and once using all available scores to form class means (full sample).

The following predictors were considered:

1. Individual CAT score (used only in analyses with restricted sample)
2. Individual CAT score squared (CAT^2) (used only in analyses with restricted sample)
3. Class mean CAT
4. School mean CAT
5. Grade (7 vs. 8)
6. Period (first observed section of a given teacher vs. second observed section).

The following conclusions were reached after performing these analyses:

1. The section period observed was not a significant predictor. (This was not surprising.) Therefore, further analyses were not done separately by period.

2. Grade level did not contribute significantly to the prediction of English achievement, and, although significant, it had only a weak relationship to math achievement (r = .06). Therefore, further analyses were not done separately by grade level.

3. Once the class mean CAT was entered as a variable to predict achievement, inclusion of a school effect did not significantly improve predictability of residual achievement scores.

4. Using the adjusted means based upon the full data set versus the data set based upon only pupils having all scores made very little difference. In both English and mathematics, the multiple R's based upon class means computed from pupils having both CAT and achievement scores differed by less than .01 from the multiple R's based upon means that included some pupils with one of the scores missing.

5. Since the initial set of process-outcome analyses to be performed was to use the class as the unit of analysis, class mean achievement was the most appropriate criterion to use, and class mean CAT was the most appropriate predictor to use as an ability covariate. Such a decision not only kept the outcome measures at the same level of aggregation as the process measures, but it also allowed use of the full sample of students' scores to compute class averages.

6. Relationships between CAT and achievement in math were high

enough to allow confidence that the achievement CAT was an effective covariable to adjust for student entering ability, but there still was room for meaningful prediction of achievement from classroom or teacher behaviors (21% of the variance was not accounted for by CAT). However, the English achievement test was more strongly predicted by CAT, leaving only 14% of the variance unaccounted for. (The process-outcome results reported in later sections are much more easily interpreted for math classes than English, a finding that may be due to this factor.)

7. Student Ratings of Teachers did indeed seem to be a separate outcome, tapping something different than the achievement test that was not predictable from the CAT scores.

Analytic Methodology

Examination of various prediction models through multiple regression techniques led us to single out class mean CAT (adjusting achievement scores from a given classroom for the average CAT for that classroom) as the control method to be used for testing additional regression models constructed to determine which of many high and low inference measures of teacher behavior were related to gain in mathematics and English achievement. The "class mean CAT" control allowed us to use all available scores and to control for school differences, tracking within schools, and grade level. In other words, once "class mean CAT" was entered into the prediction equation, these latter variables did not add to the prediction of class mean achievement.

In addition, we wished to determine the degree and direction of process-outcome relationships, to determine whether the relationships

were comparable at different levels of initial ability, and finally to determine whether the relationships were nonlinear. For the purpose of these analyses, each class section was treated separately in the analysis. To determine whether process-outcome relationships depended upon subject matter, all analyses were conducted separately for math (N = 58) and English (N = 78).

Data analyses treated each class as a distinct unit rather than pooling the two classes for each teacher, because we are restricting inferences about teacher effects to those specific to individual classes. This was considered necessary in view of marked differences between classes of the same teacher where a teacher might be effective with one group and not with another. Pooling the two classes for each teacher could mask these possible differences.

Tests of process-outcome relationships were conducted using two sets of linear regression equations for each of the potentially predictive teacher or classroom behavior variables. One of the equations (listed below under "Linear Relationships") included the degree of simple relationship of the process variable to gain and also the degree of the variable's interaction with initial student ability (Veldman and Linsley, Note 6).

Linear Relationships

The three regression equations used in this set are shown below. As indicated, each produces a squared multiple correlation coefficient, and selected comparisons of these R^2 values yield F-ratios and associated probability values that test whether particular variables improve the prediction of class mean achievement.

$$\text{Post Ach} = \text{Pre CAT} + \text{CB} + (\text{CB})\ (\text{CAT}) + E_1 \qquad R_1^2$$

$$\text{Post Ach} = \text{Pre CAT} + \text{CB} + E_2 \qquad R_2^2$$

$$\text{Post Ach} = \text{Pre CAT} + E_3 \qquad R_3^2$$

Test 1: Interaction Effect

$$F_I = \frac{(R_1^2 - R_2^2)}{(1 - R_1^2)/(N-4)} \qquad df = 1,\ (N - 4)$$

Test 2: Main Effect

$$F_s = \frac{(R_2^2 - R_3^2)}{(1 - R_2^2)/(N-3)} \qquad df = 1,\ (N - 3)$$

In these equations "POST" is the criterion achievement test given at the end of the school year, "PRE" is the CAT measure of initial ability, "CB" is the particular classroom behavior being assessed, and "E" represents errors of prediction. Each equation is solved for a set of weights that minimize the E values, thus maximizing R^2, which is an index of the amount of criterion variance associated with the predictor variables in the equation.

The R^2 associated with the first equation must equal or exceed that of the second, which must in turn equal or exceed that of the third, because each equation contains successively less information (i.e., fewer variables). The product variable in the first equation represents the interaction of initial ability and teacher or classroom behavior, and the first F-test therefore assesses whether the relationship is the same at all levels of initial ability. The second model assumes the relationship is the same at all ability

levels, and then tests whether the relationship is significantly greater than zero. Because the class pretest mean appears in all equations, initial differences between the achievement levels of the classes are "statistically controlled." For example, the second comparison asks whether the posttest is predictable from the teacher behavior beyond what is predictable from the pretest. In other language, we are asking whether holding initial ability constant, classes that were exposed to different levels of the teacher or classroom behavior, differ in their posttest achievement.

In the event that the interaction is found to be statistically significant ($p \leq .05$), expected values for the posttest are calculated for particular combinations of pretest level and classroom behavior, in order to explicate the nature of the interaction. Four combinations are presented:

Low Pre with low CB

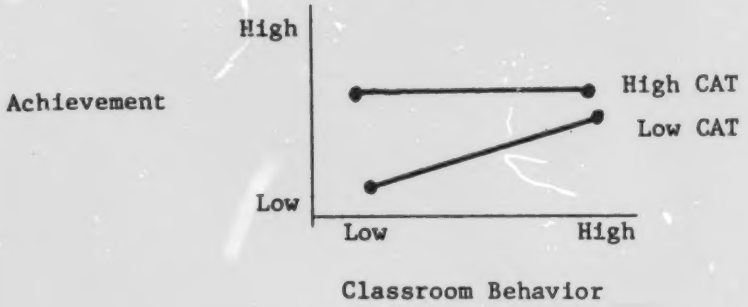
Low Pre with high CB

High Pre with low CB

High Pre with high CB

where "high" and "low" are plus and minus one standard deviation from the mean of the variable concerned. To facilitate comparisons across classroom behavior variables, these values are scaled as z scores (mean = 0, SD = 1). In the example below, we see that the behavior is positively related to gain, but that its effect is restricted to classes whose initial ability is low (low CAT). The

achievement of classes whose initial CAT scores are high do not appear to be influenced by the behavior.



The second test, which forces the implicit regression lines to be parallel, may or may not be significant, independent of any interaction effect. If both tests are significant, we still can make a general statement about the classroom behavior's effect, but with a qualification recognizing its interaction with initial ability.

In the event that only the second test is significant, we can determine the direction of the effect of the classroom behavior simply by examining the sign of the CB beta weight in the second equation.

Results

Presentation of the results of this study is complicated by the sheer number of significant relationships. For clarity, we will attempt to describe patterns that make interpretive sense and to emphasize not so much significant individual variables as the patterns that emerge from clusters of variables with similar relationships to achievement.

The accompanying three papers will deal with aspects of data from the study. The paper entitled "High-Inference Ratings of General Behaviors"will report relationships between ratings of classroom processes and the student achievement outcomes including discussion of results for subject matter and ability level of students.

The second presentation, "Low-Inference Coding of Specific Behaviors, will deal with the relationships between the coding of classroom behaviors and processes and pupil outcomes. Findings are compared and contrasted for math and English classes and also for high and low entering ability groups when the California Achievement Test interacted significantly with the classroom process variable in question.

The last paper, "Comparison of Findings for Cognitive and Affective Criteria", will include results from the second outcome measure, student ratings of teachers, and will examine the relationship between this criterion and the process variables. Relationships between this criterion and those for student achievement will also be compared.

Reference Notes

1. Evertson, C., Anderson, L., Edgar, D., Minter, M., & Brophy, J. Investigations of stability in junior high school math and English classes: The Texas Junior High School Study (R&D Report No. 4051). Austin, Texas: Research and Development Center for Teacher Education, 1977.
2. Brophy, J., & Evertson, C. Texas Teacher Effectiveness Study: Classroom coding manual (R&D Report No. 4038). Austin, Texas: Research and Development Center for Teacher Education, 1976.
3. Brophy, J., Evertson, C., Baum, M., Crawford, J., & Edgar, D. Junior High School Study: Coding manual (R&D Report No. 4014). Austin, Texas: Research and Development Center for Teacher Education, 1975.
4. Emmer, E. Classroom observation scales (R&D Report No. 4062). Austin, Texas, 1973.
5. Tikunoff, W. J., Berliner, D., & Rist, R. An ethnographic study of forty classrooms (Tech. Rep. Nr 75-10-5). San Francisco: Far West Laboratory for Educational Research and Development, 1975.
6. Veldman, D., & Linsley, T. Programs JHI and CURVE. Special purpose program using PRIME library routines to solve and compare linear models. Austin, Texas: Correlates of Effective Teaching Program, Research and Development Center for Teacher Education, 1977.

References

Brophy, J., & Evertson, C. Learning from teaching: A developmental perspective. Boston: Allyn & Bacon, 1976.

Kounin, J. Discipline and group management in classrooms. New York: Holt, Rinehart, and Winston, 1970.

Process-Outcome Relationships in the Texas Junior High School Study:

High Inference Ratings of General Behaviors

Linda M. Anderson

Carolyn M. Evertson

R&D Center for Teacher Education
The University of Texas at Austin

and

Jere E. Brophy

Michigan State University

This report is part of a four paper session entitled "Process-Outcome Relationships in the Texas Junior High School Study," which was presented at the Annual Meeting of the American Educational Research Association, Toronto, Canada, March, 1978. This paper is a general summary of part of the data from the study and should not be cited without further reference to the full report.

This study was supported in part by the National Institute of Education Contract OB-NIE-G-78-0116, The Research and Development Center for Teacher Education. The opinions expressed herein do not necessarily reflect the position or policy of the National Institute of Education and no official endorsement by that office should be inferred.

This paper will present partial data from the Junior High School Study, conducted by the Correlates of Effective Teaching Program (COET) at the Research and Development Center for Teacher Education at The University of Texas at Austin. The focus of the paper is the relationship between classroom processes (as described by six types of high inference measures) and achievement in both math and English classes in junior high schools (grades 7 and 8). It is part of a four paper package, and other papers describe the background of the study as a whole and present other data. All process-outcome data from this study will be presented in a technical report to be released this spring.

The data discussed in this paper are "high inference," in that they represent global impressions, rather than counts of discrete behaviors. The six types of high inference measures were:

1. *The Classroom Observation Scales*. Fifteen scales completed during each observation yielded information about types of teacher questioning and interaction styles. Four factor scores were derived. The scales were based on behaviors or classroom processes which were commonly included in classroom observation systems.

2. *Observers' Ratings of Teachers' Methods and Practices*. At the end of the year, the classroom observers completed 115 scales on each of the teachers. These differ from the Classroom Observation Scales in that they represent summary or overall impressions given after several hours of observations in each classroom. There were several separate ratings for each of three components: classroom

management, personal-social interactive style of the teacher, and methods of teaching academic content. Five factor scores were also created for each class on the basis of these ratings.

3. *Classroom Descriptions*. A third source of high inference data about teachers in their classrooms was a set of written classroom descriptions completed after each observation. The instructions to the observers were to describe important or most salient points of each observation period. At the end of the year, the set of descriptions for each class was read by two persons who rated them on 31 five-point scales. Therefore, each class had a score on 31 scales which represented the set of classroom descriptions. This data set differed from the rating scales described above in that these ratings were based on relatively unstructured descriptions of the most salient aspects of each classroom visit.

4. *Student Ratings of Teachers*. At the end of the year, the students were asked to rate their teachers on nine five-point scales to determine the students' opinions of the teachers' competence and personal relationships with students. Factor scores were also created from these nine rating scales. These were averaged for each class.

5. *Observer Ratings of Students*. In addition to measuring aspects of teacher behaviors and classroom processes, high inference data were also obtained on individual students. These were averaged for each class in order to gain a picture of student behaviors and characteristics most evident in that classroom. There were 26 five-point scales and four factor scores.

6. *Teacher Ratings of Students*. The teachers were also asked to rate students in their classrooms on five five-point scales. The scores were averaged to obtain a mean score to represent the teacher's perceptions of student characteristics in his or her class.

In summary, the high inference data discussed in this paper represent several different approaches to measuring aspects of classroom life which may be viewed globally.

This paper presents a summary and brief discussion of the significant results ($p \leq .05$) for the high inference variables. Results are presented separately for math and English classes.

Tables 1, 2, and 3 list the significant variables and describe the process-outcome relationships in terms of their slopes. The significant variables have been grouped to describe classroom management, personal/social/affective characteristics, and academic styles and techniques.

Summary of High Inference Data for Math Classes

When considered together, the data sets depict a fairly consistent view of effective teaching in junior high math classes, when effectiveness is defined by achievement on a test of math content.

Classroom management. Most of the measures describing management showed clear, noninteractive relationships with math achievement. In all cases, better management was positively related to achievement. This was evident for general descriptions of classroom management, such as the rating "Effectiveness of teacher's management methods," as well as for several specific descriptions of student

behaviors and classroom routines. For example, the following variables were positively related to achievement: student respect for teacher, student obedience, teacher consistency in enforcing the rules, teacher monitoring, and efficiency of transitions. Some variables showing negative relationships were: high numbers of disruptive students, classroom interruptions, and time spent in talking and "fooling around."

Personal/social/affective characteristics of the teacher and students. Variables placed under this heading included overall evaluations of the teacher's "likeability," personal characteristics of the teacher and the students, and descriptions of the atmosphere of the classroom and the general tone of teacher-student interactions. Also included here are the student ratings of teachers, which describe student attitudes toward the class and teacher.

Variables which tapped overall evaluation by the observers of the teacher's "likability" and general competence showed positive relationships with math achievement, and there was some evidence that this might be most important in lower ability classes.

Teacher personal characteristics which were related positively to achievement were enthusiasm and enjoyment of teaching. Within low ability classes, teachers who were more confident and less anxious produced higher achievement. However, these characteristics were not as important in high ability classes.

Several student characteristics demonstrated relationships with achievement, when computed as the average for each class.

Classes with higher numbers of students who were uncooperative and abusive of others had lower achievement. There were interactions with ability for some variables describing students' social interactions--extroversion was negatively related to achievement within low ability classes, as was student use of profanity. However, the better the peer relations within a class, the higher achievement was for low ability classes. These data suggest that low ability classrooms in which students evidence much social activity that is negative in tone will have lower achievement, but classes in which better peer relations are evident will be more conducive to achievement. However, this was not the case in high ability classes, where there was a negative relationship with achievement for good peer relations. Perhaps social contacts have different meanings in low vs. high ability classes.

Measures describing the general tone of teacher-student interactions suggested that in higher achievement classes, the teacher reacted positively to student feelings and created a warm, supportive atmosphere. However, these results were from one set of data only, the Classroom Descriptions, and may have been only tapping the extremes of these continua. (Not every class received a score on these variables.) Similar measures taken on the total sample of classes in the Observer Ratings of Teachers did not reveal these patterns.

There were interesting interactions for variables describing student respect for the teacher and teacher credibility. These

aspects of teacher-student relationships were important for low ability classes, but not for highs. Since these are undoubtedly related to the effectiveness of classroom management, and since the management variables were important for both high and low classes, it would be expected that credibility and respect would also be important in high classes. Perhaps teacher "presence" and aplomb was more important in the low classes than the high.

The student attitudes toward their teachers and class, as reflected in the student ratings, showed that classes with more students who liked their teacher and rated him or her highly also achieved more. Therefore, teachers who were doing a better job of teaching the content were also better liked by their students. Within math classes, there was not an obvious trade-off between attitudinal and cognitive outcomes.

Academic styles and techniques. Measures considered here include descriptions of instructional formats used, the emphasis the teacher placed on academics, and certain student and teacher characteristics that can be related to academic tasks.

There were clear findings suggesting that moderate use of public class discussion was related to achievement, rather than minimizing this technique in order to provide mostly seatwork.

Use of higher cognitive level questions was also related to achievement, and two specific measures of this showed strong relationships within low ability classes, although not within high classes.

Measures describing the use of much self-paced, individualized work showed negative relationships with achievement. These findings were somewhat surprising, and may be accounted for by considering the management problems that often accompany their use. Teachers who overemphasize individual work may be spending so much time dealing with the details involved that individual students actually receive less direct teacher instruction and monitoring than they would in a setting that allowed for more public sharing of information.

Several variables indicated that teacher attitude toward academics was important in that teachers produced more achievement when their behaviors indicated that they were concerned about achievement, offered much encouragement to their students to perform well, and urged them to take responsibility for their own work. The use of class time also reflected this: classes with higher achievement had more time spent on productive work, and less time spent in "busy work."

Teacher characteristics which were important were overall positive involvement with the class and adequate preparation for teaching. Classes which included students who were more highly motivated and persistent, with fewer problems such as poor work habits and low interest, were the classes with higher achievement. Unfortunately, it is impossible to separate those student characteristics which were independent of the teacher, and those which were reflective of the atmosphere established by the teacher in that class.

In general, effective math teachers were portrayed by the data as being in control of their classes with smooth management systems

in effect; being relatively warm and accepting, although the data do not support this as strongly as other characteristics; being very serious about academics; and running classes which included a moderate amount of public discussion and highly efficient use of instructional time to perform relevant tasks.

These results make sense, considering the content of junior high math courses. The students are learning a systematic body of knowledge which they are unlikely to encounter outside the classroom. They are also unlikely to be able to structure that knowledge for themselves in order to learn it. Therefore, the role of the math teacher is clearly defined: it is to serve as the primary source of information about the learning of math and to structure the learning experience in a way that will facilitate retention of that systematic body of knowledge. Apparently, and not surprisingly, math classes in which this was done well were those which were run in a task-oriented, business-like manner, but also in which the students could function comfortably without undue concerns about peer problems or relations with the teacher.

Summary of High Inference Data for English Classes

Unfortunately, the data for English classes did not hang together as clearly as they did for math.

Classroom management. There were only four management variables which showed significant relationships with English achievement, and these did not form a coherent pattern. One variable describing overall management was positively related, although this came from the

Classroom Descriptions, which might have tapped extreme groups of excellent and poor managers; therefore, the relationships would not necessarily hold up for the entire sample. Indeed, ratings of management taken on all the teachers yielded no significant results.

Teacher use of explanations to solve behavior problems was related to achievement, which might indicate that more verbal teachers who used language to deal with problems were more effective. However, this conclusion was not supported by other high inference data. It must be concluded that these data did not shed much light on the importance of management in junior high English classes.

Personal/social/affective characteristics. There were more significant relationships found among this cluster of variables, but most of them showed interactions in which a clear relationship was only evident for low ability classes.

No overall observer ratings of the teacher's "likability" were related to achievement, and the only teacher personal characteristic to show a relationship was the rating "Teacher seems to enjoy teaching."

Student characteristics demonstrated several significant results, however. Evidence of physical maturity and athletic ability showed negative relationships with achievement, and this was usually most evident in low ability classes. For both high and low groups, classes with more students rated as having good peer relations had lower achievement. These results suggest that students whose primary

interests in class were probably not academics were not as likely to do well on the achievement test which tapped word usage, spelling, and grammar. Since very few classroom or teacher variables demonstrated relationships with achievement, it may be that English achievement was most strongly predicted by student characteristics. If this is the case, then these results make sense. Classes with more students who exhibited interest in or skill with athletics and social relations would not be expected to score as highly as classes whose students were more academically oriented.

Descriptions of the tone of teacher-student interactions did not yield significant relationships, except for one variable from the Classroom Descriptions: "Students cooperate with others and teacher." This showed a positive relationship with achievement in low ability classes, but no pattern in high classes. In the absence of other data suggesting that teacher efforts at providing such an atmosphere were important, it cannot be determined if this might also reflect student characteristics.

The student ratings yielded several significant interactions with ability. In all cases, there were negligible to slightly positive relationships with achievement in high ability classes, but negative patterns within low ability classes. That is, in low level classes, performance on the achievement test and student attitude toward the teacher and the class were inversely related. Perhaps teachers who emphasized the type of information covered on the achievement test (word usage, grammar, spelling) aroused negative feelings

among their students in low level classes, who might find such content frustrating and less reinforced outside the classroom.

Academic style and techniques. There were few significant results for these variables for English achievement. There was a negative relationship for the total sample for teachers' adjusting learning pace to student learning rates. This was somewhat unexpected, and may represent the dangers of an extreme use of such techniques to the point that little progress is made and the teacher holds very low expectations. Perhaps related to this was the rating of "Teacher consistently gives feedback on assigned work," which showed a negative relationship within low ability classes and no relationship for highs. Again, extreme use of such techniques might slow down the pace of the class so much that little content is covered, especially the systematic type of content emphasized on the English test. The frequency of homework was positively related to achievement in low ability classes. This variable might represent the opposite of the other two significant variables: more use of homework might indicate a more demanding, faster paced approach to teaching English content. However, these interpretations are tentative, since they are not supported by other high inference data for English classes.

Not unexpectedly, student characteristics showed some relationship to achievement. Student motivation and teachers' ratings of student academic performance demonstrated interactions with ability, such that there were positive relationships with achievement in low level classes, but no relationship within high classes.

These findings provide further support for the idea that student characteristics are more important determinants of performance on a test of English grammar and word usage than were the classroom processes and teacher behaviors observed. This is probably due to the content of English classes, in that grammar was a component of the course, but not the only part, or even the major part. If grammar is not emphasized, students' performance will be more easily predicted by their personal concerns, their language learning background, and the attitudes they have toward its study.

It may be that systematic learning of English usage could be facilitated by the same type of class as was described for effective math teaching. However, in this study, no English classes were set up to totally emphasize word usage, grammar, and spelling. Therefore, this hypothesis was not tested. Instead, the high inference data revealed that within the classes observed, student achievement on the English test was more dependent on entering student characteristics than on classroom processes determined by the teacher.

Table 1

Significant Relationships Between High Inference Measures of Classroom Management and Adjusted Achievement Scores

		Interactions (Math)		Main Effects		Interactions (English)	
		Low	High	Math	English	Low	High
GENERAL DESCRIPTIONS:							
02003	Effectiveness of teacher's management methods			+	0		
02065	Factor 1: Effective organization, control			+	0		
11022	Teacher emphasizes quietness, orderliness, and good behavior			0	+		
11030	Teacher is in control of the class and maintains order			+	0		
STUDENT BEHAVIORS:							
02008	Student obedience to teacher			+	0		
03002	Student is obedient, does not defy the teacher			+	0		
11029	Students respect the teacher			+	0		
03022	Student has behavior problems, disrupts class frequently			-	0		
03020	Student is continually talking to neighbors	-	0	I	0		
SPECIFIC INDICATIONS OF MANAGEMENT:							
02004	Crowdedness of classroom			0	+		
02014	Consistency of enforcement of rules			+	0		
11004	When teacher makes a threat, it is followed out			+	0		

Table 1--Continued, Page 2

		Interactions (Math)		Main Effects		Interactions (English)	
		Low	High	Math	English	Low	High
SPECIFIC INDICATIONS OF MANAGEMENT: (cont.)							
02021	Monitoring of class			+	0		
02022	Efficiency of transition during the class period			+	0		
02017	Teacher uses explanations to solve behavior problems			0	+		
02010	Classroom interruptions			-	0		
02015	Teacher grants requests to go to restroom or water fountain			-	0		
02012	Teacher has seating arrangement	0	+	I	I	0	-
02013	Frequency of seating arrangement changes			-	0		
11031	Time is spent in activities such as off-task talking, fooling around			-	0		

Note: Explanations for symbols: "+" indicates a positive relationship between the high inference variable and the learning criterion; "-" indicates a negative relationship; "0" indicates no significant relationship; and "I" indicates a significant interaction. Interaction data are given only when a significant interaction effect was observed. "Low" refers to low ability students; "High" refers to high ability students. The variable numbers correspond to those used in the tables in the technical report which is available from the R&D Center. The first two digits of the number denote the source of data: 01 (Classroom Observation Scales), 02 (Observer Ratings of Teachers), 03 (Observer Ratings of Students), 04 (Student Ratings of Teachers), 05 (Teacher Ratings of Students), and 11 (Classroom Descriptions).

Table 2

Significant Relationships Between High Inference Measures of Personal, Social, and Affective Characteristics of the Teacher and Students and Adjusted Achievement Scores

		Interactions (Math)		Main Effects		Interactions (English)	
		Low	High	Math	English	Low	High
OVERALL EVALUATION OF TEACHER:							
02064	Child, if 7th or 8th grader, would choose this teacher			+	0		
02069	Factor 5: Teacher competence, confidence	+	0	I	0		
11032	Observers' overall positive evaluation of teacher			+	0		
PERSONAL CHARACTERISTICS OF THE TEACHER:							
02028	Teacher confidence level	+	0	I	0		
02027	Teacher anxiety	-	0	I	0		
02029	Teacher enthusiasm			+	0		
11015	Teacher seems to enjoy teaching			+	+		
PERSONAL CHARACTERISTICS OF THE STUDENTS:							
03001	Student is very outgoing or extroverted	-	0	I	I	-	0
03016	Student has good peer relations	+	-	I	-		
03018	Student has chip on shoulder and engages in physical or verbal abuse			-	0		
03021	Student lacks cooperativeness, shows no desire to work with others, etc.			-	0		

Table 2--Continued, Page 2

		Interactions (Math)		Main Effects		Interactions (English)	
		Low	High	Math	English	Low	High
PERSONAL CHARACTERISTICS OF THE STUDENTS: *(cont.)*							
03013	Student is physically mature			0	-		
03023	Student has athletic ability			0	I	-	0
03028	Factor 3: Physical, athletic development			0	I	-	0
03024	Student uses profanity often	-	0	I	-		
TEACHER-STUDENT RELATIONS:							
02030	Student respect for the teacher	+	0	I	0		
02034	Teacher credibility	+	0	I	0		
03017	Student has good relationship with teacher			+	0		
11001	Teacher reacts positively to student feelings			+	0		
11005	Warm, family-like feeling to classroom; positive feelings among class			+	0		
11006	Students cooperate with others and teacher			+	I	+	-
11010	Teacher divides time and attention among all students			+	0		
11020	Teacher acknowledges students' feelings, both positive and negative			+	0		
STUDENT RATINGS OF THE TEACHER:							
04001	Students think the teacher knows the subject well			+	0		
04002	Students think the teacher is always well prepared and organized			0	I	-	0

Table 2--Continued, Page 3

		Interactions (Math)		Math Effects		Interactions (English)	
		Low	High	Math	English	Low	High
STUDENT RATINGS OF THE TEACHER: (cont.)							
04003	Students think the teacher enjoys teaching			+	I	-	0
04004	Students think teacher has interest in knowing class as well as teaching			0	I	-	0
04005	Students feel comfortable asking questions or asking for help			0	I	-	0
04006	Students feel comfortable about going to the teacher with a personal problem			0	I	-	0
04007	Students feel they have learned a great deal in the class			+	0		
04009	Students would ask for this teacher again			+	0		
04010	Factor 1: Generalized liking of the teacher			+	I	-	0
04011	Factor 2: (females only) Student view of teacher competency			+	0		
04012	Factor 3: (males only) Student view of teacher competency			+	0		
04013	Factor 4: (females only) Favorable teacher/student relationships			+	I	-	0
04014	Factor 5: (males only) Favorable teacher/student relationships			+	I	-	0

Note: Explanations for symbols: "+" indicates a positive relationship between the high inference variable and the learning criterion; "-" indicates a negative relationship; "0" indicates no significant relationship; and "I" indicates a significant interaction. Interaction data are given only

Table 2--Continued, Page 4

when a significant interaction effect was observed. "Low" refers to low ability students; "High" refers to high ability students. The variable numbers correspond to those used in the tables in the technical report which is available from the R&D Center. The first two digits of the number denote the source of data: 01 (Classroom Observation Scales), 02 (Observer Ratings of Teachers), 03 (Observer Ratings of Students), 04 (Student Ratings of Teachers), 05 (Teacher Ratings of Students), and 11 (Classroom Descriptions).

Table 3

Significant Relationships Between High Inference Measures of Academic Styles and Techniques and Adjusted Achievement Scores

	Interactions (Math)		Main Effects		Interactions (English)	
	Low	High	Math	English	Low	High
QUESTIONING TECHNIQUES, CLASS DISCUSSIONS:						
01002 Teacher initiated problem solving	+	0	I	0		
01007 Higher cognitive level student behavior	+	0	I	0		
01014 Higher level cognitive questions: synthesis, why questions			+	0		
02048 Time allotted for class discussion			+	0		
02059 Teacher primarily uses class discussion			+	0		
02067 Factor 3: High use of class discussion vs. seatwork			+	0		
02037 Receptiveness to student input			+	0		
02041 Teacher use of blackboard for lectures and discussions			+	0		
03015 Student participates in class			+	0		
USE OF INDIVIDUALIZED INSTRUCTIONAL TECHNIQUES:						
02039 Variety and choice in assignments			-	0		
02040 Teacher use of self-paced work			-	0		

Table 3--Continued, Page 2

		Interactions (Math)		Main Effects		Interactions (English)	
		Low	High	Math	English	Low	High
USE OF INDIVIDUALIZED INSTRUCTIONAL TECHNIQUES: (cont.)							
11014	Teacher assigns learning tasks to match individual abilities/interests			-	0		
11021	Teacher perceives student learning rates and adjusts learning pace			0	-		
EMPHASIS ON ACADEMIC WORK:							
02036	Academic encouragement given by teacher			+	0		
02056	Teacher concern for academic achievement, grades			+	0		
11009	Teacher encourages student effort; gives support for work			+	0		
11023	Teacher encourages students to take responsibility for their own work			+	0		
02053	Frequency of homework			0	I	+	0
02054	Amount of class time spent in productive work			+	0		
02063	Teacher consistently gives feedback on assigned work			0	I	-	0
11011	Teacher fills empty time with busy work			-	0		
TEACHER CHARACTERISTICS:							
01016	Factor 1: Positive affect, teacher involvement			+	0		
02050	Amount of teacher preparation			+	0		

Table 3--Continued, Page 3

		Interactions (Math)		Main Effects		Interactions (English)	
		Low	High	Math	English	Low	High
TEACHER CHARACTERISTICS: (cont.)							
02052	Teacher academic effectiveness			+	0		
STUDENT CHARACTERISTICS:							
03009	Student is highly motivated and eager			+	I	+	0
05003	High student academic performance, compared to rest of class			0	I	+	0
03004	Student has bad work habits, short attention span, unprepared to respond			-	0		
03014	Student lacks persistence			-	0		
03019	Student is irresponsible, doesn't turn in work on time, comes without supplies			-	0		
03026	Factor 1: Low achievement motivation, low interest, poor work habits			-	0		

Note: Explanations for symbols: "+" indicates a positive relationship between the high inference variable and the learning criterion; "-" indicates a negative relationship; "0" indicates no significant relationship; and "I" indicates a significant interaction. Interaction data are given only when a significant interaction effect was observed. "Low" refers to low ability students; "High" refers to high ability students. The variable numbers correspond to those used in the tables in the technical report which is available from the R&D Center. The first two digits of the number denote the source of data: 01 (Classroom Observation Scales), 02 (Observer Ratings of Teachers), 03 (Observer Ratings of Students), 04 (Student Ratings of Teachers), 05 (Teacher Ratings of Students), and 11 (Classroom Descriptions).

PROCESS-OUTCOME RELATIONSHIPS IN THE TEXAS JUNIOR HIGH SCHOOL STUDY: LOW INFERENCE CODING OF SPECIFIC BEHAVIORS[1]

Jere E. Brophy, Michigan State University

and

Carolyn M. Evertson, Research and Development Center

For Teacher Education, University of Texas at Austin

[1]This is one of a set of papers on Process-Outcome Relationships in the Texas Junior High School Study, presented at the Annual Meeting of the American Educational Research Association, Toronto, March, 1978. The remarks and accompying tables included herein were prepared specifically for oral presentation at the convention, and should not be cited. A more complete and formal presentation of the findings will be sent to individuals requesting copies of the convention presentation. This research was supported by National Institute of Education, Contract OEC 6-10-108, Research and Development Center for Teacher Education, and by Contract OB-NIE-G-78-0116, Correlates of Effective Teaching Program. Dr. Brophy also is supported by NIE Contract 400-76-0073, The Institute for Research on Teaching. The opinions expressed herein do not necessarily reflect the position or policy of the National Institute of Education, and no official endorsement by that office should be inferred.

(The data to be described are part of the Texas Junior High School Study. The overview and general design for this study will have been presented in a previous paper by Carolyn Evertson, so that this paper will move immediately to a presentation and discussion of the findings from the low inference coding of teacher-student interactions.)

As Carolyn indicated, the 58 math classes and 78 English classes were observed an average of 20 times each with a complex system for coding a great deal of information about interactions between teachers and individual students. The system was developed for this study, but it had much in common with systems used in previous work at the Texas R & D Center, in particular the one used in the Texas Teacher Effectiveness Study conducted at the 2nd and 3rd grade levels. This is because the present study at the Junior High School level was intended in part as a replication. Many of the findings from the 2nd and 3rd grade levels were surprising, at the time. They contrasted with certain other lines of process-outcome research, especially the work of Flanders and his associates, which had been conducted mostly at the junior high and high school levels. Methodological differences could explain the contrasting findings, but we believed that the contrasts were real and based on differences in the nature of learners and of the teaching-learning situation between the early elementary grades and the junior high school grades. Consequently, we wanted to test these assumptions by replicating our earlier work at the junior high school level, adapting our data collection instruments to accomodate differences in what occurs in classrooms at these different levels, but retaining as much as possible in order to maximize the elements that would be replicated across the two studies.

As with the earlier work, the present study yielded process-outcome relationships that apply across the range of student ability, as well as interactions that

suggest different relationships in low vs. high ability classes. In addition to this distinction, the present study involved comparisons of relationships in math vs. English classes. The patterns of relationships observed in these two subject matter areas were quite different, so that they will be discussed separately. Throughout this paper, I will be discussing relationships between classroom behavior variables and student learning criterion scores computed by adjusting scores on tests given at the end of the school year by the class average on the California Achievement Test administered previously. All main effects or interactions discussed were significant at the .05 level or better in the multiple regression analyses described earlier. (See Tables 1 and 2 for summaries of the rate and proportion data, respectively.)

MATH CLASSES

The findings from the 58 math classes replicate many of the main trends seen in 2nd and 3rd grade, and they also show some expected contrasts on variables believed to interact strongly with grade level. The most pervasive set of correlates of learning in math classes was a pattern indicating better achievement in classes that featured frequent public verbal interaction. In these classes, teachers maximized interacting with the whole class, asking questions and giving feedback, relative to other activities such as private interactions with individual students or interactions devoted to dealing with procedural matters, socializing, or student behavior problems. The more effective teachers spent most of their time teaching, introducing content and interacting with the students, and less time having the students work individually. Students, especially those in the lower ability classes, appeared to benefit from this greater exposure to instruction in the whole class setting.

Within these public response situations, greater learning was associated with

a fast paced academic focus. This meant high rates of explanation (process) and fact (product) questions relative to choice or personal opinion questions, especially process questions that called for explanations of the thinking behind answers, not just the answers themselves.

The more effective teachers were better at eliciting answers from their students, whether correct or incorrect. Their students were less likely to fail to make any response at all. This is part of a larger pattern indicating greater student initiation and responsiveness in the more effective classes.

Five different types of response opportunities were coded: when the teacher moved through the class systematically, according to some kind of pattern; when the teacher called on a student before asking the question, presumably using a random or at least non-patterned method; when the teacher asked the question and then called on a volunteer, when the teacher asked the question and then called on a non-volunteer; and when a student called out an answer and was recognized by the teacher. Of these alternatives, calling on volunteers who had raised their hands was associated with student learning, across both levels of student ability. Recognizing students who had called out showed an interaction indicating that this was associated with learning positively in the high group but negatively in the low group.

These findings contrast with findings on types of response opportunities in second and third grade. There, calling on students in a predetermined patterned order was associated with learning, and calling on volunteers was not. We believe that both sets of relationships are valid and reflect important differences between early elementary school and junior high school. Young children are taught in small groups much of the time, where going around the group in a predetermined patterned order is a feasible and efficient strategy, but junior high students are taught in whole group settings where ordered turns are much less appropriate. It is more im-

portant for junior high teachers to keep moving at a good pace than it is to have public interactions with each individual student, and calling on volunteers is one way to do this.

The contrasts in findings concerning student call outs also seem related to known differences across the two settings. In the early grades, high ability classes contain students who are eager to respond, often overly eager. To keep student attention on the lesson rather than on jockeying to be called upon, teachers must suppress call outs (and parenthetically, calling on students in a predetermined patterned order is a good way to do this). By the junior high level, few if any students are still so highly motivated to "perform" for the teacher that they are eager to call out answers. Call outs usually occur in situations where the teacher invites them by addressing difficult questions to the class as a whole. Thus, at this level, the kinds of call outs that used to be problems are no longer an important factor, and certain call outs are legitimate and even helpful to the teacher (for one thing, they help maintain a rapid pace).

Lower ability classes contrast with these higher ability classes at both grade levels in regard to the meanings and functions of student call outs. In the early grades, teachers working with low ability classes have to combat student inhibition and fearfulness. Their problem is getting the students to be willing and able to respond, not suppressing their excess enthusiasm. Consequently, call outs usually are a good sign in this kind of classroom, even where their content is incorrect. Such call outs indicate that the teacher is succeeding at least in holding students' attention and getting them to respond. By junior high, though, the advantages in call outs have become less important than the disadvantages. Inhibition and fearfulness are no longer the severe problems they were in earlier grades. Also, call outs now are more likely to be undesirable in their effects, slowing down the pace of lessons and perhaps interrupting them altogether if what the student says is irrelevant, distracting, or even threatening to the teacher.

To summarize, our data indicate that the central functions of calling on students in the early grades are making sure that each student gets frequent and relatively equal opportunities to respond and get feedback during public recitation situations, regardless of student ability or eagerness. By junior high, teacher concern has shifted from maximizing interactions with individuals to conducting smoothly flowing and well paced whole class presentations. Interactions with individuals remain important, especially when the teacher wants to check with a member of the "steering group" to assess the probable depth of understanding across the group, but in general, individual response opportunities are subordinated to the objectives of the group lesson at this grade level.

The importance of maintaining an appropriate pace is also seen in several other findings. For example, in the more effective classrooms, incorrect student responses typically occurred when the teacher called on volunteers or called on a student randomly before asking the question. In contrast, the less effective teachers were more likely to create incorrect answers by calling on non-volunteers (students who had not raised their hand when given an opportunity to do so). This could indicate that the less effective teachers had difficulty in getting their students to be willing to respond, and/or that they too often risked ruining the pace of the lesson and embarrassing their students by calling on those who did not raise their hands.

Following correct answers, effective teachers typically either praised (this was significantly associated with learning only for the high group) or gave minimal feedback before going on to the next question. They did not break the flow of the lesson by changing the subject to non-academic matters (again, this was significant only for the high ability group). When students did respond but responded incorrectly, they did not try to elicit an improved response by simplifying the question or giving clues. Usually, they simply gave the answer, although two other possibilities for handling this situation showed interactions across ability groups. Calling on another student for the answer was positively associated with learning in the low

ability group, but negatively in the high group. On the other hand, integration of an incorrect answer into the flow of the discussion was associated positively with learning in the high group but negatively in the low group.

These different patterns probably result from several factors. First, the level of discourse isusually lower in low ability classes, where a series of purely factual questions is typical. Here, it often makes sense for the teacher to just move on to another student when the first one doesn't give the answer. Where there are more higher level questions, though, it is more feasible and probably more effective to integrate an incorrect answer into the discussion. Also, some of the general differences between high and low ability classes noted in the early grades still exist in residual form in junior high. The high ability classes are still relatively more competitive and eager to respond, and the low ability classes are still relatively inhibited. Thus, calling on someone else to respond when the first student has given an incorrect answer probably is more appropriate in low ability classes. It seems less likely to cause resentment or embarrassment in view of existing student expectations and classroom responding norms.

In contrast to these situations where students responded but incorrectly, teacher behavior that was effective when students said that they didn't know the answer or made no response at all was somewhat different. Here, the more effective teachers were more likely to respond either with criticism of the student or with process feedback (giving an explanation geared to make the point clear to the student). This indicates that the more effective teachers were highly demanding, making it clear to students that they expected them to pay attention and be prepared for lessons, although they also were willing to re-teach when necessary. Similar patterns were evident in our second and third grade data.

When teachers called on students prior to asking a question, students were relatively more likely to fail to answer the question correctly. Sometimes, teachers persisted in trying to get them to respond by simply repeating the question. This

behavior was associated with learning positively in the low ability classes but negatively in the high ability classes. Similar findings were noted in our earlier study. The general principle here seems to be that persistence in trying to elicit an improved response is helpful if it is likely to succeed (simply repeating the question might enable a low ability student to answer correctly the second time, because it was not heard or understood correctly at first or because the student benefits from additional time to think about it; however, high ability students usually would answer the question in the first place if they knew the answer. Also, this technique interrupts the pace of the lesson, and a rapid pace seems to be more important for high ability classes than for low ability classes).

Although calling out answers by students was positively associated with learning only in the high ability classes, student initated questions and comments during public discussions were positively associated in both ability groups. This is part of the general finding that the more effective teachers had students that were more active and initiatory in public discussions and lessons. For high ability classes only, this even was true for student questions that were irrelevant.

The more effective teachers were unlikely either to praise relevant student comments or criticize irrelevant ones. However, they were likely to give simple feedback to relevant comments or to integrate them into the discussion. In dealing with relevant questions, they were more likely to either re-direct them to the class or to integrate them into the discussion in some other way. Thus, these teachers were receptive to student initiations. They did not interrupt the lesson to praise such contributions, but they did integrate them into the flow.

These findings on integration of student contributions into the lesson, along with certain other aspects of what Flanders has called indirect instruction, contrast with our second and third grade data. Again, we believe both sets of data are valid. Students in the early grades are learning and practicing tool skills much of the time, so that individualized, teacher dominated, direct instruction is appropri-

ate and effective. However, as tool skills are mastered and students move to learning situations calling for application, and as they become more mature, sophisticated, and able to make important contributions, indirect instruction, especially the use of student ideas, becomes more feasible and probably more effective. In any case, although we found no support for use of student ideas and related techniques in the early grades, we did in junior high school.

There were no significant relationships for the sheer frequency of any of the types of teacher or student initiated private interactions. Teacher acceptance of student initiated social interactions was positively associated with learning. Also, the more effective teachers kept their private interactions with students brief, minimizing the time they spent in extended contact with individuals. For the most part, they concentrated on academics and did most of their extended instruction in group lessons, although they were friendly in responding to the students as individuals. They were more likely to respond to misbehavior with a mild correction or management statement than with threat, criticism or punishment. They were unlikely to make what Kounin has called target errors (correcting the wrong student). (There were some minor interactions in regard to correction of student misbehavior that will not be discussed here).

In summary, effective instruction in junior high math classes was marked by an academic orientation, concentration on whole group instruction rather than individualized contact, frequent public recitation and discussion with active student involvement and initiation of questions and comments, maintenance of a rapid pace by calling mostly on volunteers and minimizing lengthy interruptions to deal with the needs of individual students on the spot, and in general, a stress on eliciting and reinforcing high quality responses to questions designed to move the class along at a good pace. This stereotype fits the high ability classes much better than the low ability classes, though. Both the level of demand and the level of discourse was a little lower in the low ability classes (appropriately so), and effective teachers

in those settings spent more time dealing with individuals, especially attempting to elicit improved responses. There was much support for aspects of what Flanders has called indirect teaching, especially for the eliciting and use of student ideas, although again, more so in the high ability classes and only within the larger context of a strong academic and demanding orientation.

ENGLISH CLASSES

There were relatively few significant findings for the 78 English classes, and most of those that did appear involved interactions with student ability level rather than main effects. Further, the data do not hang together well to form interpretable patterns. Possible reasons for this will be discussed after a brief look at the findings that did appear.

The data on types of interactions did not suggest any particular type of classroom that was positively associated with learning in English. The frequency of social interactions between the teachers and students was positively associated with learning in the low ability classes, and the frequency of interruptions caused by student misbehavior was negatively associated at both levels. However, the pattern observed in math classes indicating that achievement was associated with academically focused public recitation and discussion did not appear in the English classes; nor was any other pattern generally associated with learning. Thus, all we can say is that, in English, learning was facilitated if the students were orderly and the teacher did not have difficulty controlling them.

The data on level of question and quality of student responses did not yield any important findings, either. The sheer frequency of two of the four types of questions did correlate positively with learning gains in English, but there were no significant relationships for any of the proportion variables expressing each type of question as percentage of the total questions asked. Thus, no particular type of question appeared to be more important than the other types. Also, there

were no significant main effects for types of student answers: correct, incorrect, don't know, or no response. However, a complex pattern of findings from measures of "don't know" responses occurring in certain kinds of response opportunities or following certain kinds of questions indicated that these "don't know" responses were positively associated with learning in the low ability group. Apparently, in low ability classes, student willingness to say "I don't know" rather than make no response at all is a positive indicator. The data from high ability classes are mixed and contradictory, however.

Significant interactions involving types of response opportunities also indicated findings for the low ability classes only. High percentages of student volunteering to respond was associated with good learning in these classes, but high percentages of calling out answers was not. Despite this, learning was higher in both low and high ability classes when call outs were followed either by praise from the teacher or by the teacher giving the answer (as opposed to calling on another student for the answer, simplifying the question, criticizing, or giving the extended feedback). It is difficult to know what if anything to make of these findings, except to note that they support the importance of maintaining a good pace.

There are a number of significant relationships for teacher feedback variables in combination with types of response opportunities and types of questions. But many of them conflict with one another, and there is no pattern of positive findings. An exception is praise of correct responses, a variable that was positively associated with learning at both ability levels. This contrasts with our second and third grade findings, where praise of correct answers was either uncorrelated or negatively correlated with learning, depending on the measure used, but it supports the earlier findings of Flanders and others from the junior high level. Once again, we think the differences are valid.

First, there is reason to believe that teacher praise of good answers is more genuine and reinforcing at junior high. For one thing, it is much less frequent

than in the early grades. Also, at higher grades there are many more complex and difficult questions, and thus more opportunities for students to be praised for truly noteworthy accomplishments. This makes it more likely that teachers' praise will be genuine when it does occur.

In general, our data indicate that fostering student talk, using student ideas, and praising student accomplishments are aspects of effective teaching at the junior high level, supporting the work of Flanders and others. Note, however, that all of this is in the context of high quality student talk about academic content, a situation that makes use of student ideas and praise of student accomplishment much more feasible, as well as effective, than they would be otherwise.

Also, the praise involved here is praise of accomplishment - correct answers to questions. Other types of praise, such as praise of good conduct, were not associated with learning in English classes, and one type, praise of relevant student co[illegible]rs, was negatively associated. So was criticism of irrelevant comments.

The rates of student and teacher initiated private contacts did not show main effects either, but there were some interactions. The frequency of student initiated social interactions was positively associated with learning in the low ability classes but unrelated in the high classes. Meanwhile, teacher initiated work interactions were positively associated with learning in the high classes but unrelated in the low classes.

The percentage of teacher initiated work interactions that involved criticism of the student was negatively associated with learning at both levels. The less effective English teachers were more critical of their students, and/or less successful in causing them to do acceptable work.

High rates of misbehavior were associated with poor learning, as expected. However, there were a variety of unexpected relationships involving measures of teacher response to misbehavior. For example, target errors related positively to learning in both groups, and timing errors did, too, in the low group. Also, mild rather than

severe responses to minor misbehavior were negatively related to learning in the low group. These relationships for responses to misbehavior contradict both our expectations and our previous findings, as well as those of Kounin, and we have no explanation for them.

Finally, the percentage of total private interactions that were student rather than teacher initiated was positively associated with learning for both groups. This pattern was also seen at second and third grade, where it was better if students came to the teacher than if the teacher went to the students about private matters.

In summary, the data from the English classes are disappointing. There were few significant relationships, and the ones that did appear were not very enlightening. Sensible patterns that hang together and indicate positive correlates of learning were almost absent, and many findings were puzzling and contradictory. Thus, our low inference coding system did not yield much information about correlates of student learning in junior high English classes (although some of the other methods used in the study as a whole did yield more useful results).

Why did this occur? We believe that there are several reasons. First, there is less uniformity of curriculum and instruction at the junior high level than in the early grades, and a smaller percentage of the curricular objectives held in common can be included on standardized tests. This means that a test surveying mastery of math or English in junior high is less likely to be "fair" or unbiased as an evaluation instrument. We expected, and unfortunately found, that this model of process-outcome research, featuring year-long data collection and using adjusted scores on an end-of-the-year achievement test as the criterion, is most fruitful in the early grades.

Within the present study, the same trends can be seen in comparing math and English. Math is much closer to the situation in the early grades, in that there is still a strong emphasis on skill practice and relative homogeneity of curriculum and

instructional methods across classes and schools within the same grade. There is much more variation across English classes, thus watering down the content validity of English achievement tests.

In addition, the problems of trying to cover uncontrolled naturalistic variation across many classes and schools with a single, relatively brief, test were exaggerated by the nature of the sample in the present study. We included almost all of the eligible math and English classes in 9 of the 11 junior high schools in the city. These 9 included 2 inner-city and primarily minority schools, as well as 3 others serving unusually homogeneous upper middle class populations. Despite a degree of busing for desegregation purposes, there were strong school effects in the distributions of student ability and achievement test scores. The scores themselves extended along a great range, distributed more horizontally than normally. As a result, the use of the students' California Achievement Test scores as co-variables for adjusting their scores on our achievement tests had the effect of removing even more criterion variance than is usual in these situations. This was especially the case for English classes, where some 85% of variance was accounted for by CAT scores (compared to about 79% for math). With so much of the variance in English achievement already accounted for by the co-variable, it is likely that most of the remaining variance was unreliable. This further decreased our chances for finding a rich pattern of relationships.

In any case, whatever the reasons, analysis of our low inference coding data did not yield much information about correlates of student learning in junior high English classes.

TABLE 1. Relationships Between Classroom Process Rates (Frequency Variables) and Adjusted Achievement Scores[1]

	Interactions: Math		Main Effects		Interactions: English	
	Low	High	Math	English	Low	High
Total dyadic interactions			0	0		
Total response opportunities	+	0	I	0		
Total procedural contacts			0	0		
Total social contacts			0	I	+	0
Total behavioral contacts			0	0		
Process questions			+	+		
Product questions			+	0		
Choice questions			0	+		
Personal questions			0	0		
Preselect, patterned response opportunities			0	0		
Preselect, non-patterned response opportunities			+	0		
Non-volunteer response opportunities			0	0		
Volunteer response opportunities			+	0		
Call out response opportunities			0	0		
Correct answers			+	0		
Incorrect answers			+	0		
Don't know			0	0		
No response			0	0		
Sustaining after incorrect answers			0	0		
Sustaining after don't know or no response			0	0		
Total sustaining (includes sustaining after correct answers)			+	0		
Student initiated questions			+	0		
Student initiated comments			+	0		
Total student initiations, public and private			+	0		
Student initiated private interactions			0	0		
Teacher initiated private interactions			0	0		

Table 1. cont'd.

	Interactions: Math		Math Effects		Interactions: English	
	Low	High	Math	English	Low	High
Student initiated work interactions			0	0		
Student initiated procedural interactions			0	0		
Student initiated personal interactions			0	0		
Student initiated social interactions			0	I	+	0
Teacher initiated work interactions			0	I	0	+
Teacher initiated procedural interactions			0	0		
Teacher initiated personal interactions			0	0		
Teacher initiated social interactions			0	0		
Typical misbehaviors			0	0		
Serious misbehaviors			0	−		
Negative responses to student initiations			0	0		
Criticism or punishment of misbehaviors			−	0		
Time in individual seatwork			−	0		
Time lecturing or demonstrating	0	+	I	0		
Time testing			0	+		
16 other time variables			0	0		
Peer tutoring			+	I	0	−

[1]Explanations for symbols: "+" indicates a positive relationship between the frequency variable and the learning criterion; "−" indicates a negative relationship; "0" indicates no significant relationship; and "I" indicates a significant interaction. Interaction data are given only when a significant interaction effect was observed. "Low" refers to low ability students; 'High" refers to high ability students.

TABLE 2. Relationships Between Classroom Process Proportion Scores (Relative Frequency Variables) and Adjusted Achievement Scores[1]

	Interactions: Math		Main Effects		Interactions: English	
	Low	High	Math	English	Low	High
Private interactions/private & public interactions			−	0		
Behavioral interactions/total interactions			0	−		
Process questions/total response opportunities			+	0		
Product questions/total response opportunities			−	0		
Choice questions/total response opportunities			0	0		
Opinion questions/total response opportunities			0	0		
Volunteer/total response opportunities			+	I	+	0
Call out/total response opportunities	−	+	I	I	−	0
Other types of response opportunities/total			0	0		
No response/total response opportunities			−	0		
Other types of answers (correct, incorrect, don't know)/ total			0	0		
% incorrect answers followed by simplifying the question			−	0		
% incorrect answers followed by criticism			0	I	+	0
% don't know and no response followed by criticism			+	0		
% don't know and no response followed by process feedback			+	0		
% incorrect answers followed by calling on someone else	+	−	I	0		
% incorrect answers followed by integration into the discussion	−	+	I	0		
% correct answers followed by praise	0	+	I	+		
% correct answers followed by process feedback			0	I	0	+
% correct answers followed by non-academic questions	0	−	I	0		
% incorrect answers that were preselect, nonpatterned response opportunities			+	0		
% incorrect answers that were volunteer response opportunities			+	0		
% incorrect answers that were nonvolunteer response opportunities			−	0		

	Interactions: Math		Main Effects		Interactions: English	
	Low	High	Math	English	Low	High
% volunteer response opportunities leading to don't know responses			0	I	+	−
% process questions leading to don't know responses			0	I	+	−
% choice questions leading to don't know responses			0	+		
% preselect patterned response opportunities leading to don't know responses			0	+		
% product questions that were repeated (to the same student)			0	+		
% preselected nonpatterned response opportunities that were repeated	+	−	I	0		
% opinion questions leading to process (detailed) feedback			0	−		
% choice questions leading to process (detailed) feedback			0	I	0	+
% call outs leading to process (detailed) feedback			0	I	−	0
% process questions leading to no feedback at all			0	I	+	0
% volunteer response opportunities leading to non-academic feedback			0	I	+	−
% preselect patterned response opportunities leading to incorrect answers			0	I	−	+
% call outs leading to correct answers			0	I	−	0
% call outs followed by praise			0	+		
% call outs followed by teacher giving the answer			0	+		
% call outs followed by calling on someone else to answer	0	−	I	0		
% call outs leading to simplifying the question			0	I	−	0
% preselect patterned response opportunities leading to teacher giving the answer			0	I	−	0
% opinion questions followed by integration into the discussion	+	−	I	0		
% nonvolunteer response opportunities followed by calling on someone else to answer			0	+		
% choice questions followed by calling on someone else to answer			0	+		
% choice questions followed by calling out			0	+		

	Interactions: Math		Main Effects		Interactions: English	
	Low	High	Math	English	Low	High
% relevant student comments praised			0	–		
% relevant student comments given feedback only			+	0		
% relevant student comments integrated into the discussion			+	0		
% irrelevant student comments criticized			0	–		
% relevant student questions redirected			+	0		
% relevant student questions integrated into the discussion			+	0		
% student questions that were irrelevant	0	+	I	0		
% irrelevant student questions given feedback only	0	+	I	0		
% relevant student questions that were called out			0	I	+	–
% relevant but called out student questions not accepted by teacher			0	I	+	0
% student initiated work interactions that were brief (vs. long)			+	0		
% teacher initiated work interactions that were brief (vs. long)			+	0		
% teacher initiated work interactions involving process (detailed) feedback	–	+	I	0		
% teacher initiated work interactions involving criticism			0	–		
Mild (vs. severe) responses to minor misbehavior	–	+	I	I	–	0
Mild (vs. severe) responses to serious misbehavior			+	0		
Timing errors in response to misbehavior	–	+	I	I	+	0
Target errors in response to misbehavior			–	+		
Rate of misbehavior			0	–		
% total interactions that were student questions			+	0		
% total interactions that were private rather than public			–	0		
% student initiated social contacts that were accepted by teacher			+	0		
% total private interactions that were student initiated			0	+		
% total interactions that were private and teacher initiated			0	I	–	0

[1]Explanations for symbols: "+" indicates a positive relationship between the frequency variable and the learning criterion; "–" indicates a negative relationship; "0" indicates no significant relationship; and "I" indicates a significant interaction. Interaction data are given only when a significant interaction effect was observed. "Low" refers to low ability students; "High" refers to high ability students.

Process-Outcome Relationships in the Texas Junior High School Study: Comparison of Findings for Cognitive and Affective Criteria

Carolyn M. Evertson

R&D Center for Teacher Education
The University of Texas at Austin

and

Jere E. Brophy

Michigan State University

This report is part of a four paper session entitled "Process-Outcome Relationships in the Texas Junior High School Study," which was presented at the Annual Meeting of the American Educational Research Association, Toronto, Canada, March, 1978. This paper is a general summary of part of the data from the study and should not be cited without further reference to the full report.

This study was supported in part by the National Institute of Education Contract OB-NIE-G-78-0116, The Research and Development Center for Teacher Education. Dr. Brophy is also supported by NIE Contract 400-76-0073, The Institute for Research on Teaching. The opinions expressed herein do not necessarily reflect the position or policy of the National Institute of Education, and no official endorsement by that office should be inferred.

At the end of the year, prior to taking their achievement tests, students filled out a nine-item rating scale of overall instruction. They were asked to rate their teachers on the amount they felt they learned in class, whether they felt their teachers were well prepared, whether they would sign up for the teacher again if allowed to, etc. These ratings were factor analyzed and a composite score was used as a criterion for judging student attitudes toward their classes. A number of positive relationships appeared for both math and English classes. In general, the student ratings of instruction and the student achievement gains in math classes were correlated positively and significantly. A number of the same variables which appeared related to math achievement gain were also related in the same direction to student attitudes. These are listed in the attached handout.

The handout shows clusters of findings for achievement and student ratings arranged separately for math and for English classes. Variables showing positive relationships for both criteria are listed, followed by those showing negative relationships. Within these categories, the variables are grouped under general headings: general descriptions of the classrooms and the teacher, student variables (both behavioral and attitudinal), classroom management and organization, teaching techniques and strategies, and personal/social or stylistic characteristics of teachers.

The first set of measures will be those for mathematics classes showing positive relationships for both criteria; those showing

negative relationships for both criteria; those showing positive and negative relationships for the attitude measures, but none for achievement; and those showing positive and negative relationships for achievement, but none for attitude. These are listed for both math and English classes. The numbers in the handout correspond to variable numbers used in the final report (Evertson, Anderson, & Brophy, Note 1).

Positive relationships for both math student ratings and achievement suggest that students may benefit from and may also like effective management, teacher leadership in the enforcement of classroom rules and organization, a generally task oriented atmosphere, and a large proportion of time spent in class discussion and in whole class teaching. Certain personal and social teacher characteristics are also correlated for both criteria, those such as acceptance of student feelings, teacher enthusiasm, and academic encouragement. Certain student characteristics stand out, such as student obedience, participation in class discussions, student creation of public contacts, volunteering of information, and having positive attitudes and respect for the teacher.

Few negative relationships appeared for both criteria, but those that did occur were: classroom interruptions, variety and choice in assignments, teacher use of self-paced work, student lack of persistence, and proportion of serious misbehaviors for which teacher made a target error and proportion of private dyadic contacts. These few variables suggest that students neither benefit from, nor particularly like,

classes which are disruptive, or where serious misbehaviors can occur. The findings for variety and choice in assignments, and self-paced work suggest that possibly teachers' attempts to individualize in low achieving classes is not working well, or that lack of general teacher control and organization can lead to boredom and disinterest. The findings for private teacher/student contacts suggest that the teachers' attempts to contact students individually may be a control measure teachers use when observers are in the room, or that it may be an ineffective method of instructing students.

Positive relationships for student ratings, but no relationships for achievement suggest that students' attitudes are better in classes where there is opportunity for public responding to questions, for class discussion, large amount of teacher feedback to student answers, the nurturance of student affective skills, high levels of confidence, patience in correcting errors, and teacher affection for students. In short, measures showing positive relationships for ratings tend to be those which support a high level of interpersonal interactions between students and their teacher.

The following variables were negatively related for student ratings but showed no relationships for achievement: negative affect, teacher anxiety, high proportions of disruptive behavior, and tendencies for teachers to allow misbehaviors to continue before stopping them. Student variables included high proportions of student call outs, leaving the class without permission, high proportions of student created private contacts with the teacher, and high proportions of teacher private contacts with students.

These measures suggest that classes where there is relatively little gain and where there are low ratings on student attitudes, also are somewhat disruptive, the teacher is not in control, and there is a high proportion of misbehaviors. In addition, teachers tend to manifest high anxiety, show nervousness, and have a higher proportion of negative interactions with the students.

The following measures show positive relationships for achievement, but none for student ratings: high proportions of thought and fact-type questions, student volunteering and attempts on the part of the teacher to get students to respond, criticism of nonresponsiveness, incorrect answers given process (or explanatory feedback), and acceptance of students' public questions and comments. Overall, the picture is one of high task orientation, student obedience, cooperativeness, even to the point of tutoring others. These classes appear to be characterized by an emphasis on academic work and student responsiveness.

Negative relationships for academic achievement but none for student ratings were found for the following measures: time spent in off-task behavior such as talking; high proportion of teacher errors in dealing with misbehaviors, such as allowing the misbehavior to continue too long, or reprimanding the wrong student; high rates of student misbehavior; and teacher/student contacts which were long. Certain student characteristics and behaviors were also evident, and those included ratings of the students as irresponsible, disruptive, disinterested and poorly motivated, and in general unresponsive.

Students were also viewed by observers as lacking cooperation, engaging in physical and verbal abuse, and having short attention spans. Filling empty time with "busy work" also appears to be ineffective.

As it has been stated before, the findings for math hang together and show patterns which make intuitive sense. Distinctions between what appears effective for student ratings and what appears effective for academic achievement lie largely in the affective area. Student attitudes seem to be affected by the amount of interpersonal warmth and competence that teachers demonstrate, whereas academic achievement is affected by a more businesslike, organized, but not cold, teacher approach. Teachers who have both high ratings and high achievement in their classes could be those who are able to blend appropriate amounts of interpersonal skill, warmth, and interest in students with the more task-like demands of delivering mathematics content.

English Classes

English classes are more difficult to characterize. No variables were positively related for both academic achievement and student ratings across ability groups, nor were any variables negatively related for achievement and student ratings across ability groups. Instead, there is a small and somewhat confusing set of findings for these classes. In general, positive relationships for student ratings make more sense than do those for student achievement. Yet,

there are fewer positive findings than negative findings, making interpretation more difficult.

Variables which showed positive relationships for student ratings in English classes suggest that student attitudes are facilitated by a warm, family-like feeling in the class; attractiveness of the room; a high proportion of private teacher/student contacts; the use of self-paced work (a contrast to the findings for math where this variable was negatively related to gain); and use of audio-visual aids, lectures and discussions. Teacher personal characteristics which were positively related to student attitudes were: enthusiasm, nurturance of affective skills, patience, high level of affection, and ability to deal with student personal problems.

Negative relationships for student ratings of teachers in English classes are: negative affect, passive pupil behavior, teacher anxiety, high proportions of incorrect answers, high proportions of student questions and comments which were called out, and high rates of irrelevant student questions and comments which were called out. There are a large number of teaching techniques and strategies which show negative relationships. Many of these were related to criticism of student answers after public response opportunities. Failure to give feedback after student answers was also negatively related. However, the measure showing total response opportunities which the teacher sustained is also negatively related. In general, findings for English are conflicting, and some of the relationships are more readily interpretable than others.

The following variables were positively related for academic achievement, but showed no relationships to student attitudes. There are even fewer relationships for this set. However, those that do appear indicate that teacher praise of correct answers and encouragement of student responding through praise and appropriate pacing by calling on other students, if the original selected student did not answer, is positively related. Also, long teacher work contacts with students are facilitative (this relationship is reversed for math classes). Management and organization variables indicating that the teacher emphasized quietness, orderliness, and good behavior and used explanations to solve behavior problems were also positively associated.

The positive relationship for the rating, "crowdedness of the classroom," is not readily interpretable. One suggestion is that, because of a general policy within most of the junior high schools to allow students to transfer in and out of various classes, the more effective teachers also happen to have larger classes. In addition, as the city's population has shifted toward the suburbs, the newer junior high schools in the middle and upper middle SES levels have become more crowded, while the inner-city classrooms have become less so. It is also likely that at least some of the more effective English teachers have moved to these suburban schools. Under these circumstances, the relationship of crowdedness to gain is not surprising.

Minutes of testing time are also positively related and could bear a direct relationship to student test taking skills. Students

who are practiced at taking tests in the subject matter area may also be more adept at doing well on the criterion material.

Finally, those variables which were negatively related to achievement but showed no relationships for student ratings were not too surprising. These classes were characterized by misbehaviors involving students leaving class, baiting the teacher, sassing and defying, and other serious misbehaviors. There were also tendencies for teachers to overreact to student misbehaviors. High proportions of student initiated irrelevant comments criticized, and high proportions of relevant comments praised were also negatively associated, suggesting attempts on the part of teachers to condition proper student classroom behavior.

There was no correlation between student attitudes and achievement for English classes. High ability groups tended to rate their teachers positively as did low ability groups. Analysis of the observers' classroom descriptions suggested that for the low ability classes less time was spent in learning of the systematic rules of English usage and spelling. The English classes showing low achievement but high ratings of their teachers spent less time in the structured activities of grammar and spelling, punctuation, and other facets of grammar structure. Of necessity, the English tests did not cover some of the broader verbal communication skills which some teachers may have attempted to teach. The test included a number of items of a specific content nature. In addition, the correlation between the achievement test developed by us and given at the end of the year was extremely high, leaving only a small amount of variance to be accounted for by

classroom measures (CAT (pre) and posttest r = .89). In addition to this restricted variance, several of the classroom measures taken in English classes, such as classroom observation scales, showed narrow ranges and variance in comparison with those used in math classes.

Examination of the coder descriptions of English classes which scored lower on achievement but which had high ratings indicated that in these classes English teachers used a variety of methods to reach students and few of these involved practice with the criterion material. There was extensive use of film strips, games, story reading, literature discussion, and in some of these classes discipline was a problem. One possible suggestion is that in order to keep control of the class English teachers may have sought more stimulating and entertaining activities, and that few of these were directly related to grammar instruction and other rule-governed learning.

These findings suggest that some caution should be exercised in any attempt to use student ratings in teacher evaluations. It is apparent from our data that a number of influences may impinge on student attitudes, particularly in varying ability groups. There is also the suggestion that the traditional end-of-year achievement measures may not be appropriate for capturing the kinds of learning experiences contained in many English classes. This is due partly to the varied nature and methods used in conducting classes in contexts similar to English. In contrast, math instruction and information is likely to be learned only in school related settings. It is more likely that examination of the classroom behaviors and tasks found in math classes may show a direct relationship to the attainment of those concepts.

Reference Note

1. Evertson, C., Anderson, L., & Brophy, J. Texas Junior High School Study: Final report of process-outcome relationships (R&D Report No. 4061). Austin, Texas: Research and Development Center for Teacher Education, 1978.

Summary Tables describing

Process-Outcome Relationships in the Texas Junior High School Study:

Comparisons of Findings for Cognitive and Affective Criteria

MATH

The following variables showed positive relationships for both ability groups for student ratings of teachers and for academic achievement.

GENERAL DESCRIPTIONS

Coder would choose this teacher himself (2064)

Observers' overall evaluation of teacher (11032)

Teacher academic effectiveness (2052)

STUDENT VARIABLES (Behavior)

Student obedience to teacher (2008)

Student participates well in class (3015)

Proportion of:

Student created work contacts given brief feedback (9281)

Student created public contacts (9397)

Rate of:

Student initiated questions (15214)

Student initiated comments called out (15224)

Public and private work contacts sought by student (15400)

STUDENT VARIABLES (Characteristics and Attitudes)

Student is highly motivated and eager (3009)

Students think teacher knows subject well (4001)

Students think teacher enjoys teaching (4003)

Students feel they have learned a great deal in class (4007)

Students would ask for teacher again next year (4009)

Factor 1: Generalized liking of teacher (4010)

Students respect the teacher (11029)

CLASSROOM MANAGEMENT AND ORGANIZATION

Teacher has effective management methods (2003)

Consistency of enforcement of classroom rules (2014)

Monitoring of the classroom (2021)

Efficiency of transitions during the class period (2022)

Time alloted for class discussion (2048)

Amount of teacher preparation (2050)

Amount of class time spent in productive work (2054)

Factor 1: Effective teacher organization and control of the classroom (2065)

Teacher is in control of the class and maintains order (11030)

TEACHING TECHNIQUES/STRATEGIES

High level cognitive questions: synthesis or why questions (1014)

Teacher use of blackboard for lectures and discussions (2041)

Teacher primarily uses class discussion (2059)

Factor 3: High use of class discussion (vs. seatwork) (2067)

TEACHING TECHNIQUES/STRATEGIES, continued

Teacher encourages student effort, gives support for work (9293)

Proportion of:

Answers to product questions which teacher praised (9080)

Teacher initiated work contacts which involved brief feedback (9293)

Rate of:

Preselecting students in random order for answering questions (15010)

Selecting volunteers for answering questions (15021)

Process questions receiving praise (15079)

Product questions receiving praise (15080)

Non-volunteers which are praised (15142)

Volunteers which are praised (15143)

Academic praise (15395)

PERSONAL-SOCIAL CHARACTERISTICS /TEACHING STYLE

Factor 1: Positive affect, teacher involvement (1016)

Teacher enthusiasm (2029)

Academic encouragement given by teacher (2036)

Receptiveness to student input (2037)

Teacher concern for academic achievement in terms of grades (2056)

Teacher reacts positively to student feelings (11001)

Teacher acknowledges both positive and negative student feelings (11020)

Warm, family-like feeling to classroom, positive feelings in class (11005)

Teacher seems to enjoy teaching (11015)

MATH

The following variables showed negative relationships for both ability groups for student ratings of teachers and for academic achievement.

Classroom interruptions (2010)

Variety and choice in assignments (2039)

Teacher use of self-paced work (2040)

Student lack of persistence (3014)

Proportion of:

Serious misbehaviors for which teacher made a target error (9375)

Private teacher/student contacts (9391)

The following variables showed positive relationships for both ability groups for student ratings of teachers but no significant relationships for academic achievement.

STUDENT VARIABLES (Behavior)

Higher cognitive level student behavior (1007)

Students are constantly being attended to by teacher (3005)

High rate of student responding to teacher questions

STUDENT VARIABLES (Characteristics and Attitudes)

Student eagerness for response opportunities (2047)

Student respect for teacher (2030)

CLASSROOM CLIMATE/ENVIRONMENT VARIABLES

Attractiveness of room (2002)

Rate of teacher/student contacts (15392)

TEACHING TECHNIQUES/STRATEGIES

Teacher initiated problem solving (1002)

Teacher presentation of academic information (1004)

Convergent evaluative interactions (teacher probes for right answer) (1009)

Teacher task orientation (1010)

Clarity of teacher presentations (1011)

Random questioning, memory questions, fact related questions (1013)

Questions with application to students' personal lives (1015)

Quantity of directions; overly explicit (2009)

Teacher deals effectively with student personal problems (2031)

Nurturance of student affective skills (2038)

Teacher makes productive use of own mistakes (2045)

Teacher goes to students during seatwork (2046)

Frequency of homework (2053)

Teacher primarily lectures (2057)

Teacher primarily uses class discussions (2059)

Teacher actively listens to student in reading, reciting, etc. (11002)

Teacher seems to know content of subject matter (11016)

Proportion of:

Call out students whom teacher gave no feedback (9179)

Teacher initiated contacts which were work related (9284)

Response opportunities in which teacher praised (9382)

Private work contacts (9395)

Teacher initiated public contacts (excluding behavior) (9398)

TEACHING TECHNIQUES/STRATEGIES, continued

Rate of:

Brief teacher afforded work contacts (15273)

Minutes in drill (15372)

Public response opportunities (15393)

Sustaining feedback in academic response opportunities context (15399)

Reinforcing teacher/student contacts (15409)

PERSONAL-SOCIAL CHARACTERISTICS/TEACHING STYLE

Positive affect (teacher) (1006)

Teacher enthusiasm (1012)

Teacher patience when correcting errors (2001)

Average level of teacher affection (2023)

Teacher confidence level (2028)

Teacher credibility (2034)

CLASSROOM MANAGEMENT AND ORGANIZATION

Teacher consistently plans sufficient work for class (2062)

Teacher adjusts instruction schedule to be flexible (11012)

Teacher seems prepared for class (11027)

Rate of:

Minutes in lecture demonstration (15370)

Minutes in discussion (15371)

MATH

The following variables showed negative relationships for both ability groups for student ratings of teachers but no significant relationships for academic achievement.

STUDENT VARIABLES (Behavior)

Proportion of:

Choice questions which students answered incorrectly (9052)

High proportion of student call outs

Student initiated questions called out and given feedback (9223)

Student initiated irrelevant questions called out and not accepted (9228)

Student verbal aggression (9303)

Student leaving class without permission (9305)

Student created private contacts (9387)

STUDENT VARIABLES (Characteristics and Attitudes)

Factor 3: Physical, athletic development (3028)

CLASSROOM CLIMATE/ENVIRONMENT VARIABLES

Negative affect (teacher and students) (1005)

Teacher awareness of coder (2033)

Proportion of:

Response opportunities which students answered by calling out (9013)

Private non-academic contacts (9396)

Rate of:

Total procedural contacts (15401)

CLASSROOM MANAGEMENT AND ORGANIZATION

Length of time after bell for class to begin (2016)

Proportion of:

Student created contacts related to classroom procedure (9268)

Teacher initiated procedural contacts (9296)

Disruptive misbehaviors which teacher criticized (9340)

Mild misbehaviors for which teacher made a timing error (9372)

Rate of:

Teacher underreactions to such misbehaviors as leaving class, tardiness (15315,15330)

TEACHING TECHNIQUES/STRATEGIES

Teacher primarily assigns seatwork (2058)

Proportion of:

Teacher created private contacts (9388)

Rate of:

Teacher afforded work contacts (15274)

Teacher afforded procedural contacts (15276)

Minutes spent in special activities (15373)

Rate of:

Minutes in advance organizers (15374)

Minutes in individual self-paced instruction (15376)

PERSONAL-SOCIAL CHARACTERISTICS/TEACHING STYLE

Teacher anxiety (2027)

Teacher berates or puts down student in front of others (11003)

The following variables showed postiive relationships for both ability groups for academic achievement but no significant relationships for student ratings of teachers.

STUDENT VARIABLES (Behavior)

Proportion of:

Student initiated relevant questions which were redirected (9235)

Student initiated relevant comments called out and given feedback (9245)

Student initiated relevant comments called out and integrated into discussion (9247)

Student created work contacts involving brief teacher feedback (9271)

Student created work contacts given feedback (9274)

Teacher/student contacts which were student initiated questions (9385)

Rate of:

Call outs which were praised (15144)

Volunteers given process feedback (15183)

Student initiated questions and comments (15201)

STUDENT VARIABLES (Characteristics and Attitudes)

Student is obedient, does not defy teacher (3002)

Student has good relationship with teacher (3017)

Student cooperate with others and teacher (11006)

CLASSROOM CLIMATE/ENVIRONMENT VARIABLES

Number of:

Peer tutoring situations (15381)

CLASSROOM MANAGEMENT AND ORGANIZATION

When a teacher makes a threat, it is carried out (11004)

Proportion of:

Mild teacher responses to student misbehaviors (9313)

TEACHING TECHNIQUES/STRATEGIES

Proportion of:

Process questions (9001); Product questions (9012)

Don't know/no response answers which teacher criticized (9041)

Don't know/no response answers after which teacher gave process feedback (9281)

Non-volunteers whom teacher gave process feedback (9182)

Correct and incorrect answers given by volunteers (9203, 9208)

Teacher divides time and attention among all students (11010)

Teacher encourages students to take responsibility for their own work (11023)

Rate of:

Wrong answers given sustaining feedback (15033)

Don't know/no response answers which teacher criticized (15041)

The following variables showed negative relationships for academic achievement but no significant relationships for student ratings of teachers.

STUDENT VARIABLES (Behavior)

Student is irresponsible, doesn't turn in work on time, comes without supplies (3019)

Student has behavior problems, disrupts class frequently, etc. (3022)

Factor 1: Low achievement motivation, low interest, poor work habits (3026)

Proportion of:

Answers which were no response (9022)

Product questions to which students gave no response (9057)

Incorrect answers given by non-volunteers (9207)

Student physical aggression which teacher criticized (9348)

Rate of:

Long student created work contacts (15252)

Student physical aggression which is criticized (15328)

STUDENT VARIABLES (Characteristics and Attitudes)

Student has bad habits, short attention span, unprepared to respond (3004)

Student has chip on shoulder and engages in physical or verbal abuse (3018)

Student lacks cooperativeness, shows no desire to work with others, etc. (3021)

CLASSROOM CLIMATE/ENVIRONMENT

Time is spent in activities such as off-task talking (11031)

CLASSROOM MANAGEMENT AND ORGANIZATION

Proportion of:

Teacher errors in dealing with misbehaviors (9363)

Serious misbehaviors which teacher allowed to continue (timing error) (15389)

Behavioral criticism (15404)

TEACHING TECHNIQUES/STRATEGIES

Teacher fills empty time with "busy work" (11011)

Teacher assigns learning tasks to match individual abilities/interests (11014)

Proportion of:

Response opportunities generated by product questions (9002)

Incorrect answers after which teacher simplified the question (9032)

Student created content related contacts given long process feedback (9283)

Teacher initiated academic contacts which involved praise (9285)

Teacher initiated academic contacts which were long (9288)

Rate of:

Content related contacts involving long process feedback from teacher (15263)

Criticism of student misbehavior

Minutes in individual seatwork (15297)

There were no variables which were positively related for both ability groups for both academic achievement and student ratings of teachers.

ENGLISH

There were no variables which were negatively related for both ability groups for both academic achievement and student ratings of teachers.

ENGLISH

The following variables were positively related for both ability groups for student ratings of teachers but showed no relationships to academic achievement.

GENERAL DESCRIPTIONS

Teacher academic effectiveness (2052)

Coder, if 7th or 8th grader, would choose this teacher (2064)

Observers over-all positive evaluation of the teacher (11032)

STUDENT VARIABLES (Behavior)

(Teacher ratings of students:)

High student academic performance, compared to rest of class (5003)

Student displays appropriate behavior in class (5005)

Proportion of:

Response opportunities which students answered by calling out (9013)

Correct answers (9019)

Student created private procedural contacts (9400)

STUDENT VARIABLES (Characteristics and Attitudes)

Student respect for teacher (2030)

Student has good relationship with teacher (3017)

(Teacher ratings of students:)

High student motivation, compared to rest of class (5001)

Students respect the teacher (11029)

CLASSROOM CLIMATE/ENVIRONMENT VARIABLES

Attractiveness of room (2002)

Proportion of:

Teacher/student contacts which were student created (9387)

Teacher/student contacts which were private (9391)

Positive, warm family-like feeling to classroom; positive feelings in class (11005)

CLASSROOM MANAGEMENT AND ORGANIZATION

Teacher makes productive use of own mistakes (2045)

TEACHING TECHNIQUES-STRATEGIES

Teacher presentation of academic information (1004)

Quantity of directions; overly explicit and repetitive (2009)

Teacher use of self-paced work (2040)

Teacher use of blackboard for lectures and discussions (2041)

Teacher use of audio-visual aids (2042)

Teacher primarily lectures (2057)

Teacher seems to know content of subject matter (11016)

Teacher perceives student learning rates and adjusts learning pace (11021)

Teacher encourages students to take responsibility for their own work (11023)

PERSONAL-SOCIAL CHARACTERISTICS/TEACHING STYLE

Teacher enthusiasm (1012)

Teacher patience with correcting errors (2001)

Average level of teacher affection (2023)

Teacher enthusiasm (2029)

Teacher deals effectively with student personal problems (2031)

Teacher credibility (2034)

Academic encouragement given by teacher (2036)

Receptiveness to student input (2037)

Nurturance of student affective skills (2038)

(Teacher ratings of students:)

Teacher would want the student in his/her class again (5002)

ENGLISH

The following variables were negatively related for both ability groups for student ratings of teachers but showed no relationships to academic achievement.

STUDENT VARIABLES (Behavior)

Factor 4: Students with antisocial tendencies; emotional or behavioral problems(3029)

Proportion of:

Incorrect answers (9020)

Product (fact) questions which students answered incorrectly (9051)

Student initiated relevant questions and comments called out and ignored (9221)

Student created public contacts (9397)

STUDENT VARIABLES (Behavior), continued

Rate of:

Student initiated comments (15201)

Irrelevant student questions and comments which were called out

CLASSROOM CLIMATE/ENVIRONMENT VARIABLES

Negative affect (teacher and students) (1005)

Passive pupil behavior (1008)

Teacher/student contacts which were response opportunities (9384)

CLASSROOM MANAGEMENT AND ORGANIZATION

Rate of:

Minutes spent in small groups which were not teacher controlled (15364)

TEACHING TECHNIQUES/STRATEGIES

Teacher use of drama; students read parts in plays or stories (2044)

Proportion of:

Incorrect answers which teacher criticized (9030)

Incorrect answers after which teacher asked a new question (9033)

Answers to product (fact) questions which teacher criticized (9084)

Process questions after which teacher gave no feedback (9107)

Non-volunteers who answered incorrectly (9129)

Non-volunteers whom teacher criticized (9147)

Volunteers whom teacher gave no feedback (9178)

Call out students to whom teacher gave no feedback (9179)

Response opportunities which teacher gave sustaining feedback (9215)

Response opportunities in which teacher criticized (9383)

Rate of:

Correct answers followed by nonacademic feedback (15025)

Wrong answers which were criticized (15030)

Wrong answers to which teacher gave sustaining feedback (15031)

Product (fact) questions answered incorrectly (15051)

Choice questions answered incorrectly (15052)

Product (fact) questions followed by criticism (15084)

Process questions which were given no feedback (15107)

Product (fact) questions after which teacher gave the answer (15116)

Product (fact) questions after which teacher asked another student (15120)

Non-volunteers who were criticized (15147)

Teacher afforded work contacts which were long and involved feedback (15274)

Minutes spent in small groups which were not teacher controlled (15364)

Public response opportunities (15393)

Sustaining feedback which followed wrong answers (15397)

Sustaining feedback after response opportunities (15399)

Teacher anxiety (2027)

ENGLISH

The following variables were positively related for both ability groups for academic achievement but showed no relationships to student ratings of teachers.

STUDENT VARIABLES (Behavior)

Preselected patterned turn students whc answered don't know (9055)

Rate of:

Choice questions answered correctly (15007)

CLASSROOM CLIMATE/ ENVIRONMENT VARIABLES

Crowdedness of classroom (2004)

CLASSROOM MANAGEMENT AND ORGANIZATION

Teacher uses explanations to solve behavior problems (2017)

Teacher emphasizes quietness, orderliness, and good behavior (11022)

Proportion of:

Misbehaviors in which teacher acted with target error (9363)

TEACHING TECHNIQUES-STRATEGIES

Proportion of:

Correct answers which teacher praised (9023)

Product questions after which teacher repeated the question (9088)

Choice questions after which teacher asked another student (9121)

Choice questions after which another student called out answer (9125)

Call out students whom teacher praised (9144)

Call out students whom teacher gave the answer (9189)

Non-volunteers whose turns were terminated by teacher asking another student (9192)

Response opportunities during which teacher praised student (9382)

Rate of:

Choice questions (15003)

Opinion and choice questions terminated by asking another s_udent (15121)

Volunteers whom teacher praised (15143)

Long teacher afforded work contacts (15274)

Minutes of testing time (15378)

PERSONAL-SOCIAL CHARACTERISTICS/TEACHING STYLE

Teacher seems to enjoy teaching (11015)

The following variables were negatively related for both ability groups for academic achievement but showed no relationships to student ratings of teachers.

STUDENT VARIABLES (Behavior)

Proportion of:

Student initiated irrelevant comments called out and criticized (9249)

Student initiated relevant comments which were given praise (9255)

STUDENT VARIABLES (Characteristics and Attitudes)

Student is physically mature (3013)

Student has good peer relations (3016)

CLASSROOM CLIMATE/ENVIRONMENT VARIABLES

Teacher/student contacts which were behavior related (9389)

CLASSROOM MANAGEMENT AND ORGANIZATION

Proportion of:

Misbehaviors involving students leaving class without permission (9305)

Misbehaviors involving students baiting teacher (9307)

Rate of:

Student misbehaviors which include: Sassing/defying (15282); Leaving class (15285); Baiting teacher (15287)

Teacher overreacts to student misbehavior (15388)

Serious misbehaviors (15408)

TEACHING TECHNIQUES-STRATEGIES

Teacher perceives student learning rates and adjusts learning pace (11021)

Proportion of:

Opinion questions after which teacher gave process feedback (9114)